Flóra Samu

Reputation and cooperation in social dilemma games

Institute of Communication and Sociology

Supervisors: Prof. Károly Takács, PhD

Szabolcs Számadó, PhD

Corvinus University of Budapest

Doctoral School of Sociology

Doctoral Dissertation

Reputation and cooperation in social dilemma games

Flóra Samu

Budapest, 2022

DISCLOSURE OF SUPERVISORS’ CONTRIBUTIONS

Chapter 2: Károly Takács designed and programmed the experiment; Flóra Samu and Károly Takács performed research, Flóra Samu analysed data and wrote the first version of the chapter; Flóra Samu, Károly Takács and Szabolcs Számadó wrote the final version of the chapter.

Chapter 3: Károly Takács and Flóra Samu contributed to the design; Flóra Samu conducted the experiment and performed data analysis; Flóra Samu analysed data and wrote the first version of the chapter, Flóra Samu and Károly Takács wrote the final version of the chapter.

Chapter 4: Szabolcs Számadó conceived the idea, Flóra Samu carried out the experiment, Flóra Samu analysed data and wrote the first version of the chapter; Szabolcs Számadó and Flóra Samu and Károly Takács wrote the final version of the chapter.

Chapter 5: Flóra Samu conceived the idea, analysed data, and wrote the chapter.

ACKNOWLEDGEMENT

The professional and personal help and support from colleagues, friends and my family played a major role in the completion of this dissertation.

First of all, I am grateful to my supervisors, from whom I learned the most. Károly Takács provided me exceptional opportunities. By being involved in his ERC project I was able to work on a novel topic. During this project I was enriched with the knowledge of experimental methodology. His constant confidence in my abilities was a great help, and he was not just supportive, but he actively encouraged me to be involved in international research and publish in recognized international journals. He provided me with all the support and opportunities for continuous improvement.

I would also like to thank my other supervisor, Szabolcs Számadó, who together with Károly, involved me not only to their research but their international academic network. Combining their research questions lead and will lead to very fruitful scientific outcomes. His clear way of thinking, his teaching skills and his openness to the thoughts of others is very special.

I am thankful to all members of the RECENS research group for the motivating environment they created that made the work enjoyable even in the toughest moments. Special thanks to Dorottya Kisfalusi for the professional and personal help she gave me any time when I was needed. I am also grateful to Boglárka Herke for her intense and constant support in recent years. Moreover, the great personal support I received from Srebrenka Letina is notable and memorable. Dóri, Bogi and Srebrenka were the most important sources of motivation.

I really appreciate the feedback I received from Béla Janky during my doctoral training, who is one of the few sociologists in Hungary who are familiar with the topic of my dissertation. I thank him for directing my attention to the rare sociological work in this field. I also grateful to Ariel Mitev for his thought-provoking questions to the thesis draft.

Special thanks to György Lengyel who encouraged me to start my doctoral studies. He always allocated time to give Ph.D. students useful feedback during our studies, without exception. I am also thankful to Tamás Bartus for his attempts to reform of the doctoral program to be in line with international trends and for paying attention to the opinion and needs of doctoral students during the transformation. I would like to thank him personally for the answers he gave to my methodological questions and for drawing my attention to sociological works that are parallel with my research.

Lastly, I am inexhaustibly grateful for the support from my family and friends that are not limited to the Ph.D. period. I would like to highlight their attitude and continuous support, which provided me a peaceful environment where I could concentrate on my doctoral work without any difficulty.

TABLE OF CONTENTS

“Természetesen félek. De addig, amíg a cserépkályha rendesen átmelegszik, még lesznek emberi vonásaim. Ha kint ülnék, mondjuk egy tóparti ház udvarán, valahol az Isten háta mögött, a Kárpátokban, akkor is csupán azt írhatnám, hogy egyetlen dolog tölt el csodálattal: a csillagos ég fölöttem. És ez még nagyon kevés.”

Bartis Attila: A nyugalom, Magvető, Budapest 2008[2001] p. 325

"Naturally, I am afraid. But until the tile stove heats up completely, I'll still have human features. If I were sitting somewhere outdoors, say, in the yard of a lakeshore house, somewhere in the middle of nowhere, in the Carpathians, even then I could write nothing but that the only thing that fills me with wonder is the starry sky above me. And that is indeed very little."

Bartis Attila: Tranquillity, Translated from Hungarian by Imre Goldstein, Archipelago Books, Brooklyn 2008

1. INTRODUCTION

My dissertation begins with the concluding lines of Attila Bartis's work titled Tranquillity, which choice requires an explanation. Even those readers who are unfamiliar with the work may be familiar with certain part of the quoted text, as it reflects on Immanuel Kant's famous sentence:

“Two things fill the mind with ever new and increasing admiration and awe, the more often and steadily we reflect upon them: the starry heavens above me and the moral law within me.”

Immanuel Kant: Critique of Practical Reason

As you may have noticed, Bartis, leaves out the second part of the original quote, the one about moral law. While the novel provides an insight into the path leading to this state of incompletion, the present dissertation speculates about this statement (or lack thereof). In general, in this work, I investigate the potential external forces that may trigger moral behaviour if it does not originate within. Research on morality has a long history in sociology, and intrinsic motivations are not completely ruled out from the possible explanations of moral behaviour (Gintis 2000, Camerer 2011) – in my understanding not even by the author of the cited novel. My dissertation, however, discusses only a small fraction of the knowledge accumulated in this topic, and focuses on one external factor that contribute to a wide adoption of moral behaviour in human societies. It should be noted here that internal and external motivations are not mutually exclusive, in certain theoretical framework they could coexist (Simpson and Willer 2015), but as a sociologist my equipment is suitable for an analysis of the effects of society on individuals’ action. Even if this dissertation seeks to answer whether morality is triggered by certain external drivers, I also refer theories of intrinsic impulses when some of the findings required to do so (see Chapter 1 and 3). Furthermore, I do not rule out that people – due to biological or socialization reasons (Bauer et al. 2014) – are heterogeneous with regards to internal moral behaviour (Camerer 2011). The aim of the dissertation is to explore the conditions that, beyond initial preferences, lead to moral behaviour in our society.

The dissertation is also limited with regards to the context in which external factors are examined. Using a game theoretical approach, I examine a conflict in which it is difficult to provide external structural explanations for moral behaviour (Janky 2022), because non-iterated prisoner's dilemma games (see a detailed explanation later) have a strong conflict of interest between the best decisions for the individual and for the community. This conflict highlights the narrower definition of moral behaviour that I use in this dissertation. So far, I used the concept of morality to explain the chosen motto, but I should conceptualize the subject of my research more precisely. Therefore, I will use the term 'social norms' (Bicchieri 2005) instead of 'moral norms', to emphasise the role of an external force discussed in this dissertation. In this work, I argue that behaviour is shaped by social expectations, or in other words by social norms.

A social norm is established by a learning process, and therefore it can vary from one social group to another (Bicchieri 2005, 2016, Berger and Rodkin 2012). In this regard, social expectations for both self- and other-regarding behaviour can be established in different contexts (Miller and Ratner 1998, Herrmann et al. 2008). Having said that, we are one step closer to the specific social norm that has been discussed in this dissertation. I argue that informal prescriptions of prosocial actions that consider the well-being of others (Yamagishi et al. 2013, Curry et al. 2019) can indirectly maintain favourable outcomes for the community, even when this behaviour contradicts individuals' self-interests.

Conflicts, in which actions with the highest utility for individuals leads to the most disadvantageous outcome for the group, are called social dilemmas (Dawes 1980). This study explicitly focuses on cooperation as an expected behaviour that would be beneficial for the community. My main research question is why people cooperate with each other when it is risky and when cooperation provides lower utility to individuals. I investigate social mechanisms that are the building elements of everyday cooperation. Understanding these mechanisms provides an opportunity to avoid a socially detrimental equilibrium of social dilemmas.

As social norms are the expectations of a social environment, social sanctions are inevitably attached to individuals' actions. It has to be noted here, that social sanctions are not just the main drive of cooperation (Bicchieri 2005, Janky 2022), but they also contribute to the establishment of a signalling system that can reliably signal future behaviour of the potential cooperation partner. Social sanctions contribute to a reliable signalling system because experiences are shared with others, via public (e.g., buyers evaluate sellers on an e-commerce platform) or private channels (when opinions are spread through gossip). This way, rewards and punishment applies indirectly. Indirect social sanctions – mediated by reputation assignment – are especially important when potential cooperation partners are strangers (Nowak and Sigmund 2005). Two chapters of this dissertation examine how gossip contributes to reputation building (Chapter 2 and 3).

The reliability of reputational information is questionable when there is competition for scarce resources and individuals can improve their own position by damaging the reputation of rivals. If social status is a scarce resource (e.g., when individuals' position in a status hierarchy is relative, and depends on the position of other members of the group (Paine 1967, Anderson et al., 2015)) dishonesty about rivals is incentivised by status competition[1]. The spread of false information can harm individuals' reliance on gossip, and consequently, cooperation. With this dissertation I contribute to the scientific discussion about the conditions that ensure the reliability and the maintenance of social norms and to a lesser extent, my work can be linked to the formation of social norms, even if this link remains indirect.

The dissertation consists of four studies. In the first study (Chapter 2), I examine how competition for reputation and status affect the reliability of gossip. In Chapter 3, I test two structural conditions that may play a role in the maintenance of reliable reputational signals. Chapter 4 generally proves the essential conditions for the development of informative signals in signalling game with conflict. In the last research, I try to generalize the

---

[1] Status and reputation are interchangeable terms (Cillessen and Marks 2011, Bocskor 2021), but in this dissertation, I highlight one important distinction between the two concepts. In short, the term status is used when I want to highlight its relative nature, while a good reputation can be earned by every member of the group.

experimental results by analysing networks in schools. In this study I look for further signs of the manifestation of social rewards which was artificially operationalized in Chapter 2 and 3. In the following, I provide a coherent framework for the subsequent chapters, without repeating the detailed theoretical discussion of each chapter, to demonstrate that the four empirical studies in my dissertation are closely connected.

1.1. Cooperation in social dilemma games

So far, by using the concept of cooperation I was referring to an outcome of one type of social interactions in which two (or more) people simultaneously select one among two options: cooperation and defection. These options are named after the conflict between the individual and collective outcome of the game. In such an interdependent situation, participants consider not just their own goal, but the goal of their partner (Camerer 2003, Singer and Fehr 2005). In this dissertation I analyse an interdependent situation where conflict of interest between individuals makes it impossible to reach to the collective best outcome. For this reason, we call these situations social dilemmas (Hardin 1968, Dawes 1980). I model a two-player social dilemma (Rapoport and Chammah 1970, Axelrod 1984), where outcomes encourage players to choose the option that leads to the worse outcome, even if a better outcome is available. The situation is called the Prisoner's Dilemma (PD) (Luce and Raiffa 1957).

In the original example of the PD two persons are arrested for a minor crime, but the police suspect that they have also committed a bigger crime. To persuade them to testify, officers offer them two options. They can either testify against each other or remain silent. If one of them confesses, the other person will be placed in prison for 10 years, and the traitor is free to leave. If both confess, they get 8-8 years. Mutual silence would lead to a better outcome, they would been arrested for 1 year only for the minor crime. If they have no influence on what the other person will do, they will confess to improve their own outcome. By doing so, they will hurt the group, because mutual silence would lead to fewer years in prison. Figure 1 shows the four outcomes of the game.

		ALTER	
		Confess (D)	Silent (C)
EGO	Confess (D)	-8,-8	0,-10
	Silent (C)	-10, 0	-1,-1

Figure 1 The payoff matrix of the PD. Individual's (ego) payment depends on the decision of the other player (alter). In each cell, the first number indicates ego's payoff and the second indicates alter's payoff.

In two experiments in this dissertation, participants face one-shot PDs (see specifications in Chapter 2 and 3). Partners are randomly matched with each other by the computer. In the game, players can choose from two options: cooperate with the other player who is a stranger (C) or choose selfish behaviour (D). Players are encouraged to defect, no matter what the other party decides, because individual outcomes are higher if players behave selfishly. My research seeks to answer the question under what conditions players will shift to the collectively best outcome.

Prisoner's Dilemma does not occur frequently in people's daily lives, especially not among strangers (Columbus et al. 2020). Accordingly, not many researches attempt to solve the non-iterated prisoner's dilemma among strangers (for an exception see Janky 2020). The magnitude and difficulty of resolving the conflict is motivating, but on its own is not valid justification for considering this game. I argue that it is the reputation mechanisms that resolve the situations that makes my research relevant.

In the following I discuss theories about reputational rewards and punishments that could facilitate cooperation. After that, I show how status competition hinders reputation-based solutions of the PD. Then, I discuss the conditions that contribute to a development and maintenance of a reliable reputational signals.

1.2. How does cooperation become rational?

Due to the wide variety of motivations behind cooperative decisions it is an inexhaustible research topic for more than 60 years. The fundamental nature of the problem of cooperation creates the unusual opportunity of interdisciplinarity research. The dilemma of cooperation is addressed in biology, economics, anthropology, and sociology (e.g., Nowak 2006, Axelrod 1984, Gintis et al. 2005, Olson 2012[1965], Kollock 1998). Solving social dilemma situations among strangers it is still under close examination in contemporary research and suggested solutions vary along the different contexts. In the following, I only discuss explanations that were provided to solve a non-iterated Prisoner's Dilemma among strangers, or its multiplayer version (N-person PD) where the basic conflict exists between ego and all the other players.

The explanation presented in this dissertation is a mix of previous explanations that are successful under different circumstances. I build on the role of trust (e.g., Evans and Krueger 2010), external incentives (e.g., Bénabou and Tirole 2006), reputation motives (Nowak and Sigmund 1998, 2005), social norms (Bicchieri 2005, 2016) and guidance in coordination (Guala 2016).

In a non-iterative, one-shot PD the partners have no previous experience, so they cannot reciprocate each other's behaviour (Axelrod 1984). Since players make decision at the same time, they cannot communicate in advance (Dawes 1980). Reputational information is particularly important when potential cooperation partners do not know anything about each other, and reputation store information about past behaviour and give guidance about potential future behaviour.

Previous literature suggests that the use of reputational information for conditional cooperation (Fischbacher and Gächter 2010) can facilitate cooperation (Nowak and Sigmund 1998, 2005). The risk of cooperating with an individual with good reputation is obviously lower than the risk when the partner has bad reputation, but overall cooperation is still too risky, therefore not rational. Thus, this explanation is not sufficiently answering the question why individuals would choose risky cooperation. Mathematical models indicate that this explanation, labelled as indirect reciprocity, is indeed, a difficult path to cooperation (e.g., Ohtsuki et al. 2009).

Social rewards or punishment for norm compliance can change the structure of the original dilemma (Bicchieri 2005 p. 26). The "*desire to please others*" (Bicchieri 2005 p. 23) can increase the value of social rewards of cooperation, in a way that it exceeds the benefit of defection, and cooperation decisions will be preferred over defection but only if the other player cooperates as well. In such cases, the prisoner's dilemma becomes a so-called stag hunt game (SH)[2] (Skyrms 2004). Compared to the prisoner's dilemma, stag hunt games are the most common social situations in everyday life (Columbus et al. 2020).

The game differs from the PD because SH has two states of equilibrium from which individuals will not deviate, because deviation would lead to a worse outcome. One of these provides higher payoffs, but risk-aversion leads to a less beneficial equilibrium, thus it is also a social trap. Given these two equilibria, players do best if they do what the others will do. Unlike in the PD, where defection was the dominant strategy regardless of how the other player decides, here, the optimal choice depends on the decision of the other player. This dilemma can be solved through coordination, if there is a reliable signal that tells the players how the other player behaves. If ego can reliably signal her next action, alter will adjust his, because a coordinated action maximizes his payoff. If we consider reputation as such a signal, then individuals should start building a reputation in order to achieve a better outcome. Research on social norms opens a discussion about how reputation (or generosity (Fehrler and Przepiorka 2013) or postpone gratification (Posner 2000)) become a signal of cooperative intentions. Trusting in reputational information, that coordinates decisions, is

[2] In the example, two hunters go hunting for a hare or a stag. They have the highest payoff if they hunt a stag together. The worst outcome of the game belongs to the player who hunts a stag alone, because it certainly leads to a failure, while the other player successfully hunts hares. The most risk-averse choice for both players are hare hunting, which is apparently a social trap.

essential to facilitate cooperative behaviour, therefore the development of informative signals is especially interesting under conflict of interest (see Section 1.4 in this chapter) (see Chapter 4). My main research question is close to this line of research because I'm trying to answer how reputation remains informative in a competitive environment (Chapter 2 and 3).

Other studies argue that reputation turns two-player one-shot PDs into an indefinitely iterated multi-player game, where iteration (Kandori 1992) and the structure of the social network provides solutions to the proposed dilemma (Janky 2020). In Chapter 5, I explore the relationship between prosocial behaviour and social embeddedness in one-shot games in 20 classes in a primary school, where children are probably playing a multi-player meta-game. One of the main novelties of my research is the involvement of structural explanations among the conditions that ensure the reliability of reputational information under competition (Chapter 3). My goal is to do further research in this direction (see Direction for future research in Chapter 6).

1.3. Cooperation under competition

Theoretical research claims that people compete for social status (for a review see Anderson et al. 2015) that is interrelated with prosocial reputation (Garfield et al 2021, Zizzo 2004, Snellman et al. 2019). Higher status offers more power (Milinski et al. 2002) and material benefits (Willer 2009). The importance of relative status for individuals is well known in sociology and in social psychology (Merton 1968, Festinger 1954), but contemporary research also reveals brain processes that connects rewards to relative status (Fliessbach et al. 2007).

Evolutionary biology – similarly to sociology – seeks the motivations for prosocial behaviour in the social environment. A theory of competitive helping (Roberts 1998) claims that individuals cooperate even when it costly for them, because they compete for access to better partners, that can be earned with a good reputation and good reputation can be built by prosocial actions. Empirical evidence has already been obtained for the cooperation-enhancing effect of partner selection (Barclay 2004, Barclay and Willer 2007, Sylwester and Roberts 2010, Raihani and Smith 2015). Having better partners leads to surplus in payoffs in the long run (Lave 1962). At the same time, not only qualitative but also quantitative changes (e.g., more partners) can be earned by a prosocial reputation (Lyle and Smith 2014).

Goals related to social status and partner selection are not very different if we apply a network perspective (e.g. Newcomb et al. 1993, Pál et al. 2016, Snellman et al. 2019)[3]. For instance, there are central positions in the social network of collaboration partners that may benefit those who fill those positions (Cohen 2004, Burt 1992). If everyone is striving for the most cooperative partners, and individuals have a limited capacity for how many cooperative partnerships they can maintain, some people will not be able to establish a cooperative partnership and will be placed to the periphery in the network.

My research attempts to examine whether a reputation system can be developed to serve as a guide for cooperative decisions under competition. The question is interesting because a competition for cooperative partners, or for social status, makes the individuals interested in wrecking others' reputation by spreading false information about rivals which could make reputation signals less reliable (Hess and Hagen 2006). If public information is not available, gossip[4] can be a device of private transmissions. Without reliable signals conditional cooperation is not possible and cooperation collapses due to risk avoidance. In Chapter 2 and 3, I manipulate whether competition is attached to reputation and test its effect on the reliability of reputation and cooperation.

---

[3] See studies that infer status from sociometric measures (e.g. Newcomb et al. 1993, Pál et al. 2016)

[4] The indirect transmission of reputational signals is especially important in the dilemma I portrayed before, because participants have neither direct experience nor any opportunity to observe the past behaviour of their partner.

1.4. Signaling under competition

The original concept of competitive helping suggest that the cost of altruistic helping deters selfish individuals from using any signal of cooperative intent, therefore it can reflect unobservable personality traits of individuals such as prosociality[5]. In the scientific literature - especially in biology - it is widely believed that the reliability of signals is created by unnecessarily expensive signals. This view about the conditions for informative signalling appeared in the field of sociology, long before the famous biological example, Zahavi's (1999) Handicap theory, would have been born. In Thorsten Veblen's (1899) work, high status individuals indicated their status by wasteful (conspicuous) consuming, showing that they can also afford to waste money. Later in Eric Posner's work, Law and Social Norms (2000), we also see that costly actions are used by individuals to signal their ability to delay gratification. To this day, researchers tend to ignore the fact that the condition for the reliability of signals is more complex than the wasteful nature of the signal.

Among others (Penn and Számadó 2020), in this dissertation, I am challenging the assumption that cooperation signal remains informative due to its costly nature. For the first time, conditions under which signals acquire meaning are formalized in economics (Spence 1973) and biology (Grafen 1990). They show that it is not the cost of the signal that ensure its informativeness, but the differentiated marginal outcome of costs and benefits for honest and dishonest signallers. This means that signal costs have to be calibrated in a way that deters potential non-cooperators from using the signal of cooperative intent, while co-operators should be encouraged to use it. In the classic example, education reflects productivity of job applicants because salary is calculated as a function of different cost of investment in education along productivity. In other words, the employer set salaries in a way which exceeds the cost of education for people with high productivity but lower than the cost of education for people with low productivity. Recent work argue that signals do

[5] If people are competing for scarce resources, signals which were suitable for differentiating between cooperative and selfish intentions without competition, will no longer be informative on partners' prosocial preferences, because everybody would use the beneficial signal (e.g., generosity).

not necessarily need to be costly for honest players at all, what matters is the above-mentioned difference between honest and dishonest players (Lachmann et al., 2001, Penn and Számadó 2020). Conditions for honest signalling is addressed in general in Chapter 4.

Based on this theory of differential signal cost reputational signals remain credible if the ratio of cost and benefits differs along individuals' (stable) intent for cooperation. This intention is assumed to be internally motivated by (true) altruism (McAndrew 2002). Even if the ultimate causes of informative reputational signals are outside the scope of the present dissertation (e.g., it could have developed in the absence of competition (Lewis 1969)), and proximate causes that maintains reliability is only indirectly (Chapter 4) or partially addressed (Chapter 3), I suspect that maybe it is not altruism (anymore) that is transmitted via reputation. But what else can create stable but differentiated incentives for cooperation to keep signals informative? In Chapter 3 I test alternative explanations that could maintain credible reputation. In the last chapter of this dissertation where I outline possible directions for further research, I also report some of the results of an ongoing analysis where I analyse potential structural effects that could modify the reliability of reputational signals. Maybe honest signals about cooperative intentions do not reflect true unobservable traits (true altruism) but it is social status (Willer 2009, Baldassarri and Grossman 2013) or social embeddedness (Uehara 1990) that creates incentives to act in a cooperative manner, while higher-status individuals have contributed more to the collective interest (Willer 2009).

1.5. Analytical Approach

This study is approaching the problem outlined above with the toolkit of sociology, but it also heavily builds on results of other disciplines such as anthropology, psychology, behavioural economics, and evolutionary biology. Among these disciplines, sociology has the longest history in the analysis of social interrelatedness, and sociology's strong tradition on theory building gives strengths to theory-based research in comparison to other disciplines. By harmonizing the results of different fields, I draw attention to the benefits of interdisciplinary research which opens the door for unique collaboration and mutual respect between disciplines.

The explanation for the cooperation problem that was provided by biologists is a very similar to how a sociologist would approach this question. Parallels between biology and sociology has already been raised by Durkheim in his work in which he lays down the scientific foundations of sociology (Durkheim 1982[1895]). The basis of similarity is the micro-macro separation, that has been used by both classical (Weber 2020[1920], Elias 2000[1939]) and contemporary sociologist (Coleman 1994, Hedström 2005)[6]. In this dissertation, the macro phenomena that I intended to explain by interdependent micro actions, is the emergence and maintenance of cooperation and the maintenance of the reliability of reputation in competition.

1.6. Societal relevance of the research

Despite its brevity, this paragraph is perhaps the most important message of the dissertation. By looking at the relationships between cooperation and reputation, my goal is to detect intervention points that can be used to move out conflicted situation from a harmful state of equilibrium. Nevertheless, results can be interpreted in several ways due to abstract measures used to operationalize the subject of the research. Questions may arise what I mean by harmful or desired output. I am aware that there is no clear consensus among the people on the goals and the means considered suitable for achieving. In my opinion, cooperation can be detrimental to society or to a part of society (e.g., government's informants' network after World War II in Hungary). Moreover, competition and reputation systems can have consequences that increase inequality (unfair competition as a consequence of monopolization or other kind of power accumulation).

A good example is provided by Max Weber (2020[1920]) in his famous work 'The

[6] The original approach underwent a slight modification. Macro phenomena are not explained by another social fact (macro) but by micro level actions that are embedded in a social context.

Protestant Ethic and the Spirit of Capitalism'. Weber's example illustrates how hard work became and remained the signal of trust- and creditworthiness by the spread of Calvinism. While people who were trading with money or other money-related occupations were heavily rejected earlier (Figuera and Esposito 2019), with the rise of Calvinism the society became more permissive with monetary aspirations as it was attributed to a higher will in certain Calvinists sects. Besides hard work, the signal of wealth also changed from excessive consumption to a restrained lifestyle which was better able to sign wealth because of the more permissive attitude towards capital accumulation and a use of money to more money. This process resulted in an escalation of work and a wide engagement in this 'work competition'[7].

I would like to draw attention to the fact that social utility of my results depends also on its users if it comes to implementation. Although the dissertation unfortunately depicts conflicts only in an abstract way in a strictly controlled environment, I try to help the reader to understand my motivation and interpretation by providing everyday life examples.

1.7. Research project

The dissertation is part of the EVILTONGUE project financed by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 648693). Its principal investigator is one of my supervisors, Károly Takács. The aim of the research is to examine how gossip maintains social order. The project uses different methodologies for empirical testing, including small scale experiments presented in this work. Experiments were conducted at the Corvinus University of Budapest. Ethical and data protection principles and relevant laws. Participants were informed about the aim of the research and the procedure. They were protected by anonymity and during the experiment they had the right to withdraw their consent if they wanted. Only adults were involved in the experiment. Experimental

[7] Interestingly, the two signals of wealth still exist to this day (Feltovich et al. 2002).

procedure, the recruitment of participants, and the experimental instructions have been screened and have been approved by the Ethics Review Committee of the Centre for Social Sciences at Hungarian Academy of Sciences.

From now on, I will use the plural form of the first person to indicate the help I got from my supervisors in my work so far, and the future work of possible co-authors who are likely to be involved in further publication of my results.

1.8. Outline of the thesis

My thesis consists of four studies (Chapter 2-5). Each study gives a more detailed overview of the theory and empirical findings in the discussed domain of the main research question.

In the first study (Chapter 2) we test experimentally whether competition damage the reliance on reputation in a non-iterative two-player Prisoner's Dilemma game with stranger matching. By manipulating (1) whether reputation is a scarce resource and (2) whether direct benefits are attached to it, we were able to analyse the different incentives for reputation building and its effect on cooperation at the group level, as reputation can be built through cooperative actions. One of the main contributions of this chapter is the direct measurement of private evaluations and the involvement of gossip in the process of reputation-based cooperation. Since the target of gossip was freely chosen, people were allowed to spread false gossip about individuals with high reputation when competition was severe, and they were interested in improving their own chances of reputation building by degrading the reputation of others. In this paper, in addition to answering my main hypotheses about the effect of competition on the use of reputation, conclusions can be drawn about the development of a reliable reputational system as well.

In Chapter 3, I report the results of another experiment where we test two mechanisms that could strengthen the credibility of reputational information (aka gossip). The first mechanism that could make gossip (and as a consequence reputation) more reliable is cross-checking. Participants could verify gossip from multiple sources. The second mechanism builds on the assumption that honest gossip could be a signal of cooperative intent, therefore it is expected to be more reliable if cooperation partners can exchange gossip before they play, again, a non-iterative two-player Prisoner's Dilemma. We also manipulated competition, to test whether higher confidence in gossip due to cross-checking or social bonding is able to keep reputation informative under rivalry. In this experiment, participants competed in small groups, therefore we are able to detect dishonest gossiping about the selected rivals.

Chapter 4 is seemingly unrelated to the main research question. On the contrary, this chapter is crucial to understand how reputation signals could develop and under what conditions they provide a reliably guidance to conditional decisions in a social dilemma. This study is my supervisor invested almost the same amount of work as me to prove empirically that signals (aimed at changing the behaviour of others) do not have to be costly to be informative. We modified the costs of initially meaningless signals in a so-called signalling game in different ways proposed by two theories with different projections about the development of reliable signals. We also show that a signal can even be beneficial for a signaller, it does not have to be costly at all.

In the last study we step away from the lab and analyse the dynamics of friendships and negative relations in school classrooms with social network analysis methods, to test prosociality and reputation-based partner selection and rivalry in a real-life situation. We played 3 economic games (anonymous one-shot PD was one of them) with children and we mapped friendship relations in the classes. By analysing negative relationships, we also search for the manifestation of competing social norms are related to prosociality.

In the last chapter of my dissertation, I summarize the findings and collect the limitations of my research. Lastly, I appoint possible directions for future research.

2. SCARCE AND DIRECTLY BENEFICIAL REPUTATIONS SUPPORT COOPERATION[8]

A human solution to the problem of cooperation is the maintenance of informal reputation hierarchies. Reputational information contributes to cooperation by providing guidelines about previous group-beneficial or free-rider behaviour in social dilemma interactions. How reputation information could be credible, however, remains a puzzle. We test two potential safeguards to ensure credibility: (i) reputation is a scarce resource and (ii) it is not earned for direct benefits. We test these solutions in a laboratory experiment in which participants played two-person Prisoner's Dilemma games without partner selection, could observe some other interactions, and could communicate reputational information about possible opponents to each other. Reputational information clearly influenced cooperation decisions. Although cooperation was not sustained at a high level in any of the conditions, the possibility of exchanging third-party information was able to temporarily increase the level of strategic cooperation when reputation was a scarce resource and reputational scores were directly translated into monetary benefits. We found that competition for monetary rewards or unrestricted non-monetary reputational rewards helped the reputation system to be informative. Finally, we found that high reputational scores are reinforced further as they are rewarded with positive messages, and positive gossip was leading to higher reputations.

2.1. Introduction

Cooperation is integral part of our daily life (Olson 1965, Hardin 1968). In cooperation situations, however, there is a conflict between individual and common interests (Axelrod

---

[8] This chapter was published as: Samu, F., Számadó, S. and Takács, K., 2020. Scarce and directly beneficial reputations support cooperation. Scientific reports, 10(1), pp.1-12.

1984). The most severe case is the Prisoner's Dilemma (PD) game (Luce and Raiffa 1957) in which following selfish interests is the dominant strategy that disallows the establishment of the collectively optimal cooperation outcome that is superior for every interaction partner compared to mutual defection. Over the past decades a wide range of proposals have been made how to resolve the problem of cooperation (Nowak 2006, Van Lange et al. 2014). One of the informal solutions proposed by the theory of indirect reciprocity (IR) is the establishment and maintenance of reputations that provide guidelines for selecting the right partners but also for distinctive actions towards interaction partners (Nowak and Sigmund, 1998, 2005, Righi and Takács 2018). Empirical studies confirmed that cooperation can be established through the use of reputations that trigger conditional cooperative behaviour (Wedekind and Milinski 2000, Milinski et al. 2002, Feinberg et al. 2014, Everett et al. 2015, Ge et al. 2019).

In previous empirical studies, where reputations were shown to provide an efficient solution for social dilemmas, individuals could observe the past behaviour of others and hence they had perfect and true information on who had been cooperating and who had not (Semmann et al. 2005, Seinen and Schram 2006). In large populations, however, it is not feasible to observe past decisions of potential unknown transaction partners directly and a credible summary score is not always publicly available. The mechanism that helps to access reputational information is gossip in which individuals exchange evaluative third-party information (Foster 2004). Seminal models (and reviews) of the IR paradigm operate with the assumption (Nowak and Sigmund, 1998, 2005, Ohtsuki and Iwasa 2004, 2006, 2007) that gossip needs to be reliable to ensure that information received is attended and to ensure that reputation reflect past action (Nowak 2006).

There are, however, unresolved puzzles around the reliability of gossip. On the one hand, empirical observations show that humans lie (DePaulo et al. 1996, Kashy and DePaulo 1996) and that gossip could be used to undermine the target's reputation strategically (Duffy et al 2002, Hess and Hagen 2006). Incorporating the option to send strategically dishonest messages in an IR model in fact leads to the collapse of cooperation (Számadó et al. 2016). From the perspective of strategic motivations, dishonesty could be pro self (DePaulo et al. 1996, Kashy and DePaulo 1996, Hess and Hagen 2006) or prosocial

(Shalvi and De Dreu 2014). For example, people can lie to improve their own reputation, to destroy the reputation of their competitors or to serve the interest of their group (Shalvi and De Dreu 2014, Levine and Schweitzer 2015).

On the other hand, dice-roll experiments (Fischbacher and Follmi-Heusi 2013) consistently show that people do not lie as much as expected based on the utility maximizing "homo economicus" paradigm; i.e. "they leave much on the table" (Abeler 2019). The preference of being honest was one of the main factors behind this "truth seeking" behaviour. Similarly, research on strong reciprocity (Gintis 2000, Bowles and Gintis 2004, Fehr and Fischbacher 2003) proved that negative emotional reactions to selfish behaviour can lead to altruistic punishment (Fehr and Gächter 2002, Fehr and Fischbacher 2004). Correspondingly, prosocial gossip operates with similar underlying negative emotions and it can be used to punish (and deter) selfish behaviour (Feinberg 2012). Moreover, in experiments where participants could gossip, transmitted information was very much in line with observed choice (Sommerfeld et al. 2007, Sommerfeld et al 2008, Fehr and Sutter 2019). Gossip was observed even when it implied substantial costs for the sender (Feinberg 2012). Furthermore, studies suggest that gossip does not need to be completely accurate; even with noises, it can promote trust and cooperation (Giardini and Vilone 2016, Fonseca and Peters 2018). Last but not least, spreading reputational information honestly may involve additional advantages to the sender (Willer 2009, Raihani, and Bshary 2015), such as an increased reputation for reliability. The Supplementary Table S1, in A1 summarizes the proposed explanations behind honest vs. dishonest gossip.

The processing of third-party information is an important element of a functional reputation system. The rules regulating the assignment of reputation based on information available for the individual are called social norms (Ohtsuki and Iwasa 2004, Kandori 1992, Sugden 2004). First order social norms are conditional on the previous observed action only, for instance considering co-operators good and defectors bad (Nowak and Sigmund 1998). Higher order norms take into account also the reputation of the opponent in an observed action of another player. Cooperation is difficult to be maintained by first order social norms. Certain second and third order social norms work better as they allow justified defection, i.e., the punishment of previous defectors by defection (Ohtsuki and Iwasa 2004, 2006, 2007, 2009). Social norms have mainly been analysed in models assuming unbiased and public reputations and homogenous populations (Ohtsuki and Iwasa 2004, 2006, 2007), although the investigation has been extended also to hypocritical strategies and private situations (Ohtsuki and Iwasa 2015). In an exploration of possible social norms, eight norms ("*leading eight*") has been found to be able to sustain cooperation (Ohtsuki and Iwasa 2004, 2006, 2007). It is still an open question which of these social norms could be observed in empirical situations.

In line with the literature that recognizes the importance of gossip and reputation for cooperation, we expect that where gossip is available, it will provide relevant information on partners that enable cooperation condition on the partner's reputation. The alleged relationships between gossip, reputations, and cooperation are displayed in Fig. 1.

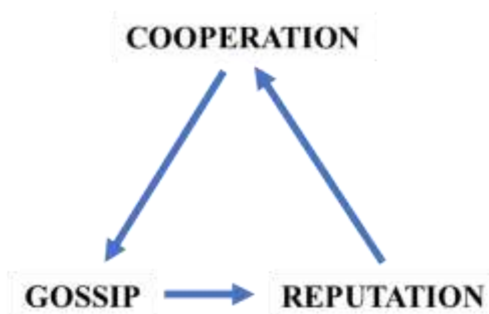


Figure 1. Schematic representation of the relationship between cooperation, gossip, and reputation.

H1 The possibility of gossip increases cooperation.

In line with the theoretical literature on reputational systems we assume that: (i) individuals use gossip to transmit their direct observations to others; (ii) reputational scores will be updated based on the information received. Therefore, we put the following mechanisms forward as sub-hypotheses:

H1a Gossip will be in line with partners' previous decisions.

H1b Reputations are updated in the direction of the valence of gossip received.

Under which conditions reputations could facilitate cooperation, however, is an open question. We propose that reputations communicated via gossip could increase cooperation if reputation is a scarce resource and hence a competitive frame is created. The theory of 'competitive altruism' (Roberts 1998, Barclay 2004, Barclay and Willer 2007) asserts that competition is the motivation for reputation-building which incorporates in the rise of pro-social behaviour. Hence, individuals compete for being more cooperative than others in order to keep up with their reputation (Barclay 2013, Henrich et al. 2015, Macfarlan et al 2013, 2015). When the highest reputation cannot be gained by everyone there is more motivation for investments in acquiring good reputation, because the individual's relative position depends on others behaviour. Therefore, we expect that individuals cooperate more in order to avoid the weakening of their relative reputation in the eye of others.

H2a Competition for scarce reputations increases cooperation.

On the other hand, relative position can be improved also by undermining the reputation of competitors (sharing negative gossip about them) (Hess and Hagen 2006, Paine 1967, Barkow 1992). So far, only competition for mates was tested empirically where romantic rivalry was taken for granted (Buss and Dedden 1990, McAndrew 2014), but there is no empirical evidence where rivalry is independent of gender. In a competitive environment for reputation, dishonest strategic gossip could also occur more likely, while there are no motivations for sending positive gossip dishonestly.

H2b Competition for scarce reputations increases negative gossiping.

As another mechanism, we propose that monetary stakes for reputation have negative effect on cooperation. Beyond partner selection, good reputation might help individuals acquiring other beneficial outcomes such as additional resources (Willer 2009), greater influence (Milinski et al. 2002), or social network benefits (Lyle and Smith 2014, Bird and Power 2015). Even if tangible incentives can foster cooperation simply because they reduce the magnitude of conflict between self-interest and the common good (Rapoport and Chammah 1965), empirical studies have shown that external incentives can reduce the motivation for reputation-building (Frey and Jegen 2001, Bowles 2008, Bravo et al. 2015, Yoeli et al.

2013) and as a result, the level of cooperation does not grow as much as we would observe in the absence of this 'crowding-out' or 'overjustification' effect (Bowles and Polania-Reyes 2012, Bénabou, R. and Tirole 2006). Either punishment as an external incentive (Deci et al. 1999, Mulder et al. 2006) or rewards (Ariely et al. 2009) can reduce motivation to achieve high reputation. The mechanism behind the reduced motivation for reputation-building supposed to be the lack of opportunity to signalling group-based motivation or commitment (Raihani and Bshary 2015, Bird and Power 2015, Johnsen and Kvaløy 2016, Gneezy et al. 2011, Smith and Bird 2000, Bird et al. 2018). We expect that if direct external incentives are linked to the reputational position, then the signal of long-term commitment or group-based motivation will be inseparable from the motives for direct benefits (Raihani and Bshary 2015, Bird and Power 2015, Smith and Bird 2000). In this case reputational signals will be less efficient, which could directly be traced in the distribution of reputations and impact cooperation as a consequence.

H3 Direct monetary stakes for reputation decrease cooperation.

We aim to show how extrinsic motivation and competition for scarce resources affect strategic reputation building and cooperation in an environment where there is a low probability of meeting with the same person again. We test whether the degree of competition influences the level of cooperation in a two-person Prisoner's Dilemma game by manipulating the scarcity of the reputational resources (H2) and further monetary incentives for reputation-building (H3). The main part of the experiment follows a 2x2 between-subject design. The scarcity of reputations is manipulated by the way participants can distribute reputation scores to others (on a scale between 0 and 100). We call treatments *abundance (A* from now on) where players can give everyone a maximum score, and *scarcity (S* from now on) where a fixed budget of scores could be distributed. Direct benefit for reputation is manipulated as reputation scores are either symbolic *(not paid, NP)* or incentivized financially *(paid well, PW)*. We expect that the impact of our manipulations will not be independent of each other. We predict that the highest cooperation level will appear under the condition when individuals are managing scarce resources, while the lowest will happen in a monetarily incentivized context where the evaluation of partners is not relative and therefore competition is less intense. The schematic representation of our experimental design are displayed in Fig. 2.

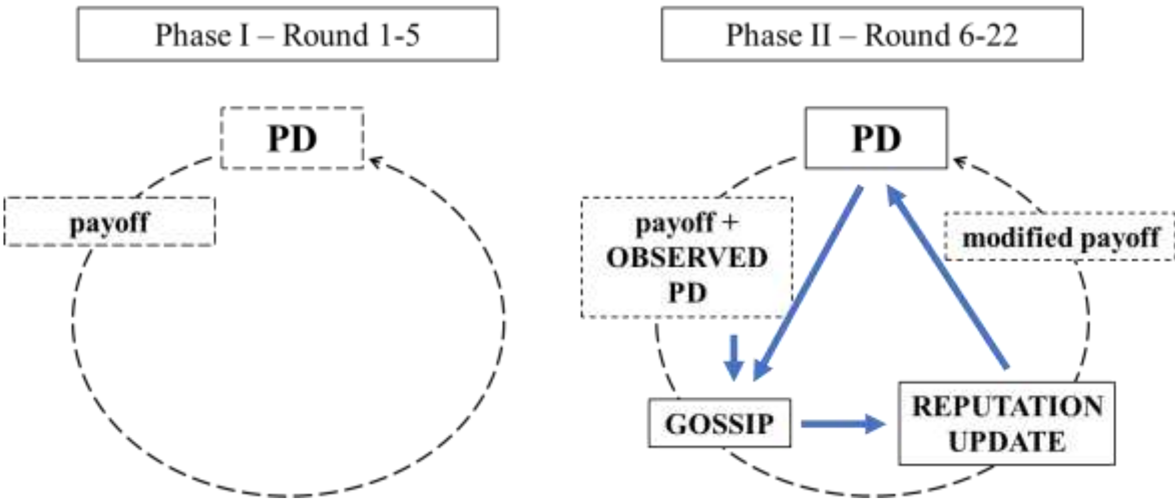


Figure 2. Description of one round in our experiment. Participants played two two-person Prisoner's Dilemma games in each round in Phase I (control) and in addition, the opportunity of gossiping and scoring others' reputation were introduced in Phase II.

2.2. Results

2.2.1. Cooperation

The introduction of gossip temporarily increased the level of cooperation in three of the four treatment conditions (A-PW: 33.8%, S-NP: 28.4%, S-PW: 32.9%) compared to the first five rounds where communication was not allowed (A-PW: 29.1%, S-NP: 21.6%, S-PW: 21.1%). Cooperation did not increase in the A-NP treatment (Round 6: 24.3%, Round 1-5: 24.5%, Fig. 3). Inspecting how decisive the changes are, we run multilevel logistic regression analysis, which revealed that the possibility of information exchange increased cooperation significantly only in the S-PW treatment (β= 0.8665 p<0.01 see Supplementary Table S1, in A2), which means only a partial confirmation of H1. Neither manipulation alone had enough positive effect to result in significantly different cooperation in the long run (see Round 7-22 effects in Supplementary Table S1, in A2). Overall, cooperation was highest in the A-PW condition (Fig. 3), because the baseline cooperation was higher.

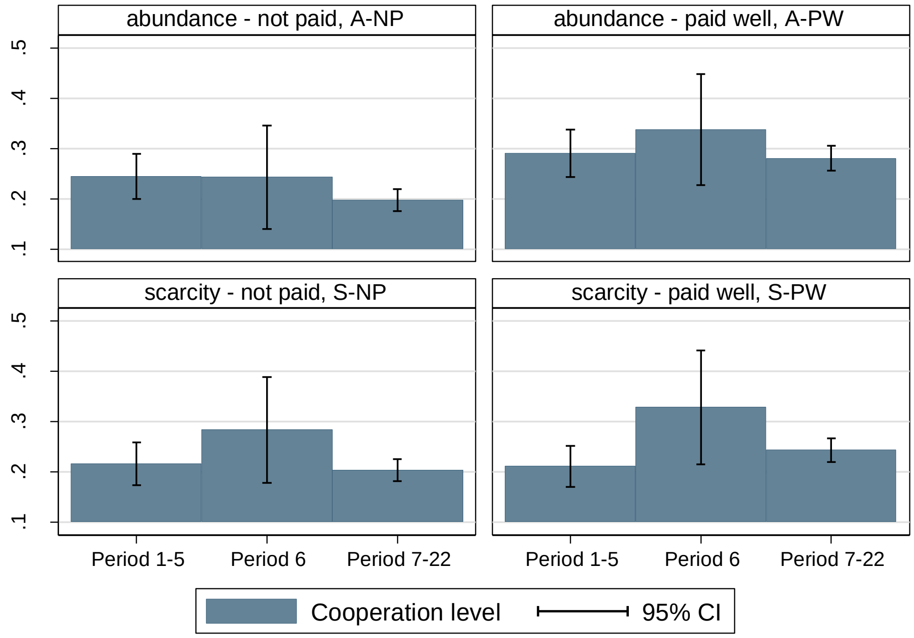


Figure 3. Average cooperation level before (Round 1-5), after (Round 7-22) and at the time when manipulations were introduced (Round 6) by treatments (see abbreviations above each bar chart: A-NP: abundance, not paid; A-PW: abundance, paid well; S-NP: scarce, not paid; S-PW: scarce, paid-well). The level of cooperation increased in three of the treatments, but the difference is significant in one case only (S-PW). Over time, in each treatment, cooperation fell back to or below the initial level, which is a typical finding in Prisoner's Dilemma experiments.

2.2.2. Gossip

In the experiment, participants could send three types of messages about the selected target: a happy, a neutral, or a sad smiley. Only 52% of the possible messages were sent by the participants (A-NP: 53.5%, A-PW: 51.8%, S-NP: 56.8%, S-PW: 45.7%). Positive gossip was more prevalent in the A-PW treatment (see Fig. 4 and ANOVA tables in Supplementary Table S2, S3, in A2).

The valence of gossip was very much in line with observed choices (H1a). A positive message was more likely to be sent if someone was cooperative (Ego cooperate - Alter cooperate: β= 3.3595 p<0.001, Ego defect - Alter cooperate: β= 2.4422 p<0.001), and a negative message if someone did not cooperate (Ego cooperate - Alter defect: β= -2.7489 p<0.001, Ego defect - Alter defect: β= -0.9184 p<0.001 see Model 2, Supplementary Table S5, in A2). Related to this effect, we found slight differences in the A-PW treatment compared to other treatment groups. On the one hand, cooperators here sent negative gossip about defectors to a smaller extent (β= 0.9974 p<0.05 see A-PW * ego cooperate - alter defect interaction effect in Model 3, Supplementary Table S5, in A2). On the other hand, gossip about observed cooperators in the A-PW treatment condition was less positive in comparison to other conditions (β= -0.8523 p<0.05 see A-PW * alter cooperate interaction effect in Model 3, Supplementary Table S5, in A2).

The model also revealed that higher reputational points increased the probability of more positive messages (β= 0.0208 p<0.001, see the effect of reputation score distributed to alter in the previous round in Model 2, Supplementary Table S5, in A2). The S-PW treatment modifies the effect of reputational position on gossip: in the presence of competition with monetary rewards strategically motivated gossip is more prevalent as the evaluation of individuals with high reputation is more negative (β= -0.0095 p<0.05 see S-PW * reputation score distributed to alter in the previous round interaction effect in Model 3, Supplementary Table S5, in A2).

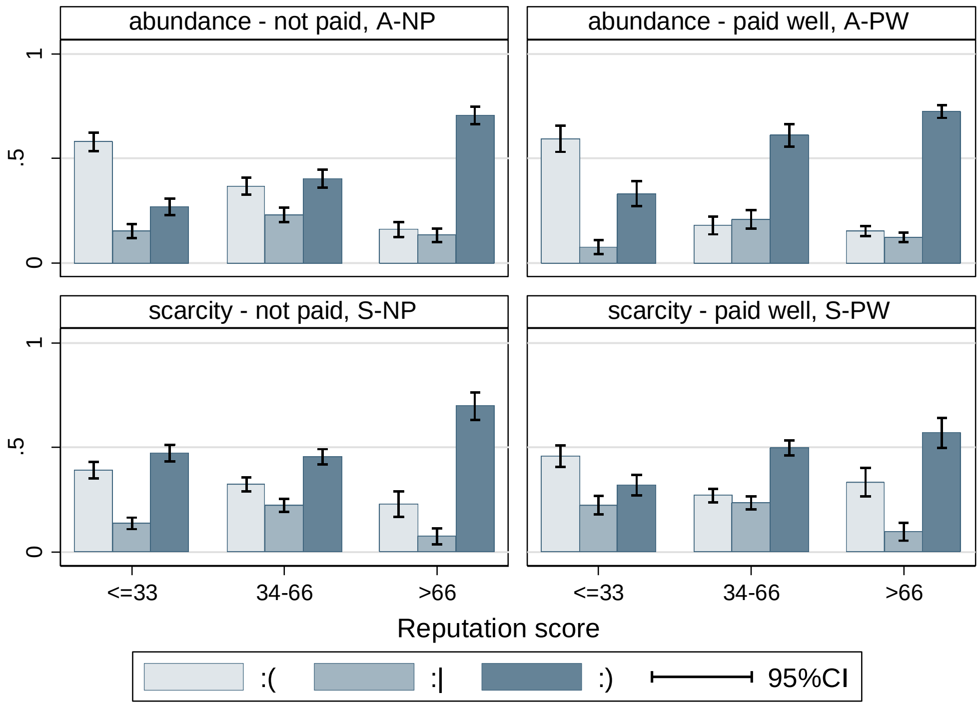


Figure 4. The distribution of gossip sent and its relation to reputation scores. Figure 4 compares the proportions (y-axis) of positive, neutral and negative gossip among individuals with low, medium and high reputation scores. Scores in players' private reputation tables were categorized into three groups (<=33, 34-66, >66). Bars shows the distribution of the valence of gossip (light blue: negative, mid blue: neutral, dark blue: positive). A breakdown by reputation score of the gossip targets shows weaker relation in S-PW between gossip targets' trust scores and the valence of gossip. Here negative gossip is more prevalent among participants with high scores, than in other treatments.

2.2.3. Reputation

The average reputation score players gave to each other is slightly lower than the initial value of 50 (M = 48.5, SD = 30.3), and the way reputation scores were distributed among participants varies in the different treatments (see Fig. 5). Higher scores are more frequent in treatment group A-PW (M = 68.4, SD = 32.0), where monetary incentives were used and participants could give high scores without lowering the scores of other players (see ANOVA tables in Supplementary Table S6, S7, in A2).

We estimated a multilevel linear regression model to explain the allocation of reputation scores. We found that the cooperative decisions of interaction partners (Ego cooperate - Alter cooperate: β= 16.0085 p<0.001, Ego defect - Alter cooperate: β= 8.7767 p<0.001) and observed players (β= 4.2662 p<0.001), as well as positive messages (β= 7.1155 p<0.001) had a significant positive effect on the allocated reputation scores, while defections (Ego cooperate - Alter defect: β= -12.3325 p<0.001, Ego defect - Alter defect:β= -2.9128 p<0.001, observed Alter defect: β= -1.3760 p<0.001) and negative messages (β= -6.9627 p<0.001) negatively influenced scores (H1b) (see Model 2, Supplementary Table S9, in A2). More messages were rewarded with higher reputation scores (see Model 2, Supplementary Table S9, in A2), but only in S-PW (Nr. of gossip sent by Alter – treatment interaction: β=0.8291 p<0.05 in Model 4, Supplementary Table S9, in A2). Looking for further differences between the treatments (see gossip - treatment interaction effect in Model 3, Supplementary Table S9, in A2) we found that in A-PW negative and neutral gossip generated a greater volume of score reduction (negative gossip: β= -5.5211 p<0.001, neutral gossip: β= -5.5822, p<0.001) and positive gossip was more powerful (β= 5.7895 p<0.001) in S-PW in comparison to other treatments.

Information from trustworthy sources might affect how much individuals rely on them. In this experiment the identity of the gossip partner was known, therefore players might have stored gossip differently when a randomly selected gossip partner had higher scores in their private reputation table. This assumption is twofold: an increase in the reputation of the gossip sender implied higher score reduction in case of negative gossip, while positive gossip causes a smaller raise if the sender is more trustworthy (see Gossip Partner's reputation - Gossip interaction effect in models for each treatment in Supplementary Table S10, in A2).

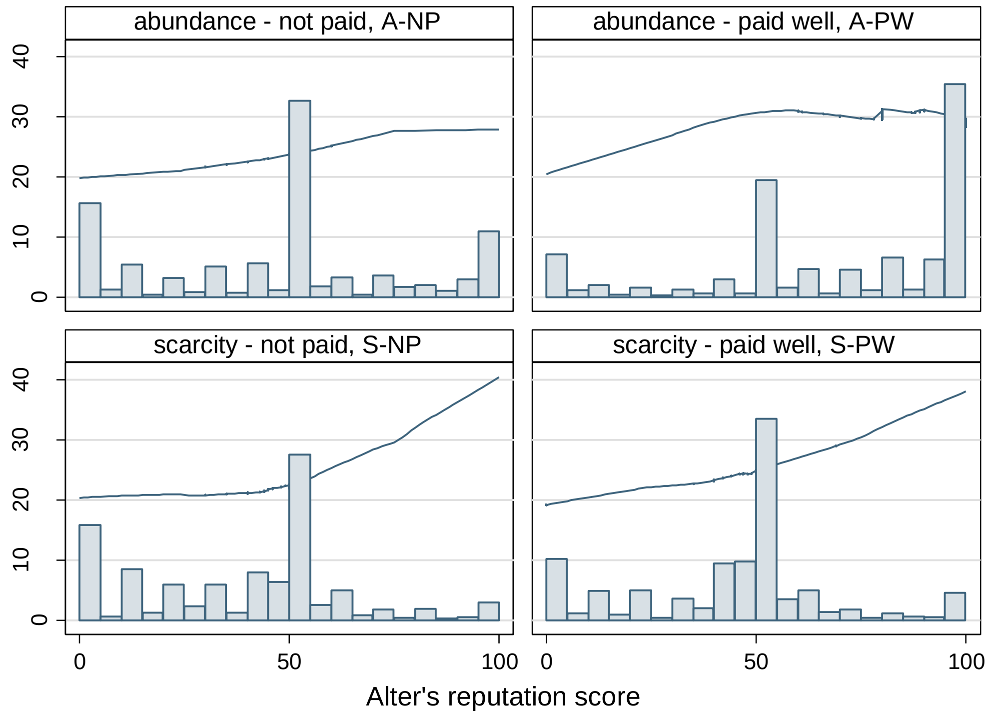


Figure 5. The distribution of reputation scores and its accuracy. Treatment groups without competition have been characterized by higher scores, and presence of lower scores was more typical in 'not paid' treatments. These trends observed along the two manipulations causes the difference in the distribution of reputation scores between the four treatments: A-PW was characterized by merely high scores and the S-NP by rather low scores. In sessions A-NP we see a wide-spread scoring in both low and high directions, while in the S-PW we observe a less extreme negative shift in the reputation score distribution. A LOWESS fitted line shows how well reputation scores reflect past behaviours. Scarcity of reputational scores better distinguishes cooperative individuals.

2.2.4. Social norms

While social norms in IR models are conceptualized as expectations on public reputation, in our experiment we are able to explore the presence of shared properties of the leading eight social norms after privately observed actions for privately assigned reputations. Multilevel mixed-effect linear regressions were used, where the dependent variable was the reputation score that has been allocated to other participants. Predicted values of updated reputations derived from the model are summarized in Table 1, Panel B (see detailed results in Model 1 in Supplementary Table S11, in A2). We do not observe changes between bad and good reputation, if we assume the default value of 50 to be the neutral point. Therefore, results should be interpreted in conjunction with predicted changes (see in Table 1, Panel C, details in Model 2 in Supplementary Table S11, in A2).

		GG		GB		BG		BB	
Panel A	C	**G maintenance of cooperation**		*		**G forgiveness**		*	
	D	**B identification of defectors**		**G justification of punishment**		**B identification of defectors**		*	
Panel B	C	**64.1**	**75.5**	61.5	69.9	**31.6**	**40**	30.6	38.3
		52.3	**55.3**	68.7	62.8	**35.3**	**44.9**	33.8	32.1
	D	**59**	**75.9**	**57.5**	**69.1**	**22.3**	**34.4**	18.6	30.2
		46.6	**48.4**	**54.2**	**55.2**	**27**	**28.8**	23.0	28.6
Panel C	C	**-1.38**	**1.18**	0.08	2.75	**13.60**	**18.78**	14.35	27.90
		-1.78	**0.27**	-3.14	-3.08	**12.94**	**16.22**	17.15	7.46
	D	**-4.36**	**-3.66**	**-4.04**	**-5.00**	**3.95**	**19.08**	4.76	16.48
		-8.42	**-4.98**	**-8.08**	**-10.08**	**2.27**	**2.57**	2.84	2.17

Table 1. Means of predicted reputation scores (Panel B) and mean of predicted changes in reputation scores (Panel C) of the focal player after the observation of a play between the focal player and his opponent in our experiment. Rows show the action of the focal player (C: cooperate, D: defect). Columns show the potential combinations of reputational scores for the focal player (first letter) and the opponent (second letter) (G: good, B: bad). Results can be compared to the table of social norms (Ohtsuki and Iwasa 2006). Expected common properties of the leading eight norms are indicated with red (good) and purple (bad) font. Each cell contains predicted reputation scores divided by treatment condition (upper-left: A-NP, upper-right: A-PW bottom-left: S-NP, bottom-right: S-PW).

There are four main features of leading eight norms (Ohtsuki and Iwasa 2006): (i) maintenance of cooperation, (ii) identification of defectors, (iii) justification of punishment and (iv) forgiveness (see Table 1, Panel A). We confirm the existence of some of these conditions of the leading eight social norms, however some others are missing. (i) The "*maintenance of cooperation*" corresponds to the characteristic that cooperation between good parties upholds good reputation (C-GG in Tables 1, Panel B and C). This feature seems to be present in our sample. (ii) "*Identification of defectors*" concerns both good and bad focal players and exists if defection against a good co-player leads to bad reputation. (iii) "*justification of punishment*" appears to be present, because despite defection good reputation remains good (D-GB in Table 1, Panel B). Changes, however, demonstrate a negative shift in reputation (D-GB in Table 1, Panel C), which contradicts the notion. (iv) Finally, "f*orgiveness"*: we cannot observe a change from bad to good reputation (C-BG in Table 1, Panel B), however, cooperative acts improve bad players' reputation (C-BG in Table 1, Panel C). The average reputation score of good focal players after such defection falls below 50 only in scarcity treatments (GG-D in Table 1, Panel B, only S-NP is significantly different β=-0.0023 $p < 0.05$, see detailed treatment differences in Model 2 in Supplementary Table S11, in A2).

All in all, the predicted reputational scores are in line with the leading eight norms in the conditions: maintenance of cooperation, justification of punishment and identification of defectors in case of bad donors, however they contradict the leading eight norms in the conditions of forgiveness and identification of defectors in case of good donors. However, the predicted change of reputational scores are in line with the leading eight norms in the conditions: forgiveness and identification of defectors in case of good donors. In other words, the absolute scores and predicted change always oppose each other. When the absolute score fits the predictions of the leading eight the predicted change does not and vice versa.

We found two outstanding effects behind these outcomes. The most important was that the focal player's action, which is considered in first order social norms significantly contributed to the focal player's reputation (β = 12.2038 $p < 0.001$ in Model 1 in Supplementary Table S11, in A2). The second obviously strong effect – as colours of Table 1 indicate – is the reputation of the focal player. These results show a clear effect of first order social norms but leave uncertainties about the functioning of higher order norms.

2.2.5. Positive assortment

A reputation system is reliable if it reflects past behaviour of others. Using aggregate statistics, we found small, but significant correlations between the level of cooperation against someone and their overall cooperativeness only in S-PW treatment (ρ= 0.38, $p<0.05$) if the whole period is considered (see Supplementary Table S12, in A2). Taking into account later periods the same correlation came into sight from Round 10 in A-NP (ρ= 0.32, $p<0.05$) and correlation in S-PW became higher (Round 10-22: 0.42, $p<0.01$). Using multilevel logistic regression models, we verified that the reputation system improved its credibility in the A-NP (β= 1.1662 $p<0.001$) and S-PW (β= 0.9393 $p<0.01$) treatments within the time available (see Alter's previous cooperative behaviour effect in Model 1, Supplementary Table S13, in A2). Partners' reputation positively influenced decision making in each treatment, but in A-PW reputation scores have lower effect on cooperation (β= -0.0278 $p<0.001$ see A-PW alter's reputation score interaction effect in Joint Model, Supplementary Table S13, in A2).

We summarized our results in Fig. 6, where the reader can follow each step of the reputation mechanism and differences between treatments. Curved arrows represent positive assortment and signs at the top of these arrows indicate the success of reputational information transmission. The association that individuals collaborate with individuals who have previously cooperated with others takes place under two conditions: external incentives with competition (S-PW) and internal rewards (cooperation) with universal access (A-NP). Successful mapping of individuals' willingness to cooperate in A-PW was hindered by the fact that cooperative third-party observations were not rewarded here with positive gossip as much as in other conditions and defectors were less punished with negative gossip. This leniency was somewhat counterbalanced as a negative gossip was followed by stronger point reduction. In S-NP, we observe reverse behaviour: negative gossip has a smaller effect on reputation. It is interesting to note that participants overrate positive gossip in S-PW. The last arrow of the triangle shows weaker reputation-based cooperation in A-PW.

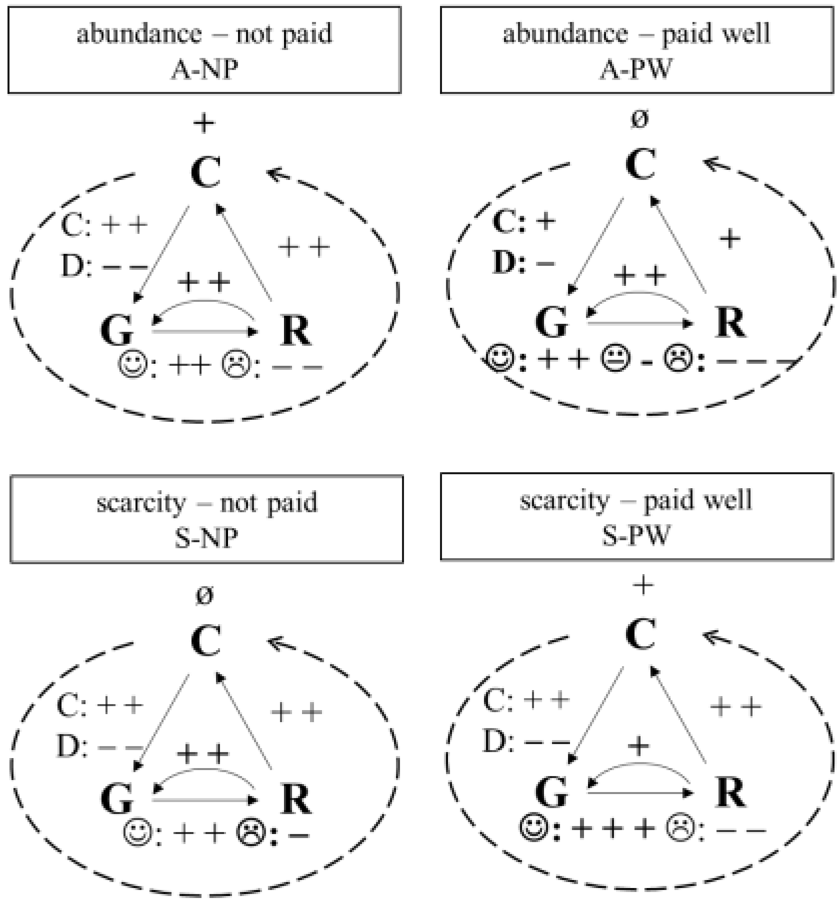


Figure 6. A summary of our results. In the figure, C, G, and R represent cooperation, gossip, and reputation respectively. The left side of the triangle shows how gossip was influenced by participant's participants' PD choices (C=cooperation, D=defection). The lower side of the triangle outlines how participants change reputation scores as a result of positive (:)) and negative (:() gossip. The right arrow represents the use of reputation scores in decision making. The relative strength of effects clarifies the distinction between treatments. The circle arrow (with its effect at the top) shows the overall accuracy of the reputation system.

2.3. Discussion

We investigated how cooperation can be sustained by private reputations formed by direct observations and gossip. We found a slight increase in the level of cooperation in one condition when the institution of reputations and gossip have been introduced (H1). Cooperation has faded over time, which is a typical feature of Social Dilemma experiments[80]. As preconditions of gossip to be effective, we expected and confirmed that gossip was in line with previous cooperation choices (H1a) and gossip received has altered the private reputations of others (H1b).

We proposed two mechanisms that could safeguard the credibility of gossip for informing the choice of cooperation. First, we investigated whether the scarcity of reputational resources (H2) with the expectation that reduced access will increase competition, could increase reputation-based cooperation. Second, we investigated whether additional monetary incentives connected to reputation would distort the credibility of reputations (H3). We found that neither the scarcity of reputational scores nor monetary incentives alone could maintain reputation on the long term. We showed, however, that at the intersections of these two manipulations, competition for scarce monetary rewards resulted in higher cooperation in the short run. To better understand our results, we discuss each treatment in detail in Supplementary Material A3.

A reputational system is reliable if it appropriately reflects the potential behaviour of others that is otherwise hidden to new partners. It functions well if it helps individuals to cooperate with those who have a higher reputation and defect against those who have lower willingness to cooperate. Our results suggest that motivation for building a proper reputation system increases if people find it easy (without competition) to credibly signal prosociality (non-monetary rewards) or if external incentives encourage everyone to participate in the competition - maybe because higher positions in the reputation hierarchy are more robust. Even though reputations have seemingly been well translated to gossip under these conditions, they did not increase strategic cooperation in a long run in the one-shot PD game with stranger matching.

We found that gossip was influenced by previous reputational scores not just by the last observed action. Since reputation scores are influenced by messages beyond actions, reputation scores could have been inflated in the informal communication process. Hence, positive gossip increased good reputations and negative gossip downgraded bad reputations. This has important implications for the whole dynamics of the development or the maintenance of reputational systems.

The analysis of social norms has revealed a similar effect: while the strongest predictor of reputational scores was the focal actor's behaviour (cooperation vs. defection), the previous reputation score also had a significant impact. In other words, there was a strong inheritance of reputational scores. Reputation updates were influenced by actions and previous reputations, but in little alignment with the "leading eight" social norms[19]. Most importantly, the reputation of the opponent had a little effect on the reputational update (see Table 1, Panel C). This has two important implications: both justified punishment and the identification of defectors might be missing from the system, i.e., defection against an opponent in good or bad state has very similar effects. While the scope of the investigation of higher order norms is limited in our experiment, yet it shows that the presence of leading eight social norms cannot be taken for granted.

The combined effect of the reputation of the focal player both on gossip and reputational updating could explain the lack of increase of cooperation in our model on the long term. Beyond the ineffectiveness of the examined reputation systems, the fact that we do not experience a larger impact of reputations on cooperation could be attributed to several other factors. Primarily, we investigated the two-person PD game with random reshuffling of partners and no publicly available information, which itself is the most severe social dilemma in which rational action is simply defection. The magnitude of conflict in the Prisoner's Dilemma game could be so strong that even a well-functioning reputation system could not increase cooperative acts (Rapoport and Chammah 1965). Unfortunately, in this study it is impossible to assess whether the magnitude of the conflict is responsible for the low level of cooperation. The PD game is an interdependent situation and hence it is not the most appropriate to test fundamental tenets of the theory of competitive altruism (Roberts 1998, Barclay 2004, Barclay and Willer 2007). Future studies could investigate if scarce reputations and direct reputational incentives could increase giving in the dictator game in the lab or in field settings. In our experiment, cooperation could have collapsed before the reputation system had been sufficiently developed, which leaves open the question of the coevolution of reputation and cooperation (Rand and Nowak 2013). It is also possible that reputation scores worked to a limited extent because they were not directly communicated to others or due to the abstract situation and scores in the experiment. Even more, participants potentially had problems to remember their earlier experiences and might have also mixed up other participants as they were identified with numbers that are harder to recall than names or faces. We should also caution about the direct correspondence of our study to the theoretical literature as we used private reputations that are realistic but contradict to the assumption of publicly available and perfect information on choices or reputations in showcased models of cooperation (Wedekind and Milinski 2000, Milinski et al 2002, Ohtsuki and Iwasa 2004, 2006, 2007). As reputations were private in our experiment, they could be used only to a limited extent for strategic reasons, and they could be linked to cooperation only through simplified gossip communication. This also limits the connection of our study to the theory of competitive altruism (Roberts 1998, Barclay 2004, Barclay and Willer 2007) as privately stored reputations cannot be used by the recipients for acquiring diverse benefits such as status, power, or access to resources, and participants could not select their interaction partners (Sylwester and Roberts 2013, Herrmann et al. 2019).

Still, our results bring us closer to understanding under which conditions reputations and gossip contribute to cooperation. Further research is needed to find out, however, under which conditions gossip is used strategically and in a dishonest way to undermine the reputation of others, and under which conditions it could be considered as altruistic punishment (Feinberg et al. 2012).

2.4. Methods

2.4.1. Participants

We investigated our hypotheses in an experimental computer laboratory with volunteer participants. The experiment was conducted at the Corvinus University of Budapest between 13-25 November 2016. In total, 160 individuals (46% female, 23.2 years old on average) participated in the experiment (male: 54.4%) in eight sessions in groups of 20. The final profit was calculated as an average payoff of 6 randomly selected rounds. In addition to the final payoff, a show-up fee (HUF 1000) has been paid to the participants. Participants earned 1822 HUF on average.

2.4.2. Procedure

The experiment has been programmed using the experimental software z-Tree (Fischbacher 2007). Participants have read the instructions on paper and on their screen after they have been randomly assigned to a computer in the lab. Subsequently, they had to fill in a quiz of understanding and when in doubt, could ask questions privately. In the experiment, participants were identified with ID numbers ranging from 1 to 20. The experiment has been divided into two phases. Phase I took place in the first five rounds (Rounds 1-5) and Phase II run for seventeen rounds (Rounds 6-22) until the end of the experiment. Subjects had no information on the total number of rounds of the experiment, which was slightly different due to time restrictions. To consider all conditions equally in the analysis we only used 22 rounds, because which corresponds to the shortest experimental session. In the second phase, players received additional instructions. In both phases, each round began with two, simultaneously played two-person Prisoner's Dilemma (PD) games. PD partners were randomly matched and IDs of the two opponents were displayed on participants' screen (please see original screens in Supplementary A5 for Phase I and A6 for Phase II). PD options were labelled with 'L' and 'R'. The cooperative decision was marked with 'L'. PD payoffs through the experiment were fixed to HUF 1500 (EUR 4.7) for mutual cooperation; HUF 500 (EUR 1.6) for mutual defection; HUF 2500 (EUR 7.8) for temptation; and 0 otherwise. Subjects had 23 seconds to decide. Results appeared on the screen after every PD game. This has completed one round in Phase I.

Rounds in Phase II were expanded with new elements of reputation and gossip. In the following, we describe these new elements in the temporal order in which they occurred on participants' screens in each round. As the first new element, on the first screen, in addition to the PD game, a 'reputation table' appeared with the IDs of all other 19 players. Next to each ID, a reputation score of 50 was displayed in Round 6. Participants were told that a value of 50 was the initial value for everyone. In later rounds, privately given reputation scores from the previous round were displayed in read-only mode. After playing the two

PD games, participants were informed about their payoffs in the PD games. On the second screen where PD results were displayed, as a second new element of Phase II, the IDs and choices of two other participants in one randomly selected PD were displayed. Hence, participants were able to observe the PD decisions of four players in total: of their own interaction partners in the PD games and of the two matched partners from a randomly selected game.

On the third screen, the next novelty of Phase II was introduced. Participants could send a maximum of four gossip to a randomly selected gossip partner (receiver), whose ID has appeared on their screen. Participants could enter up to four IDs of other participants (targets) in empty boxes on their screen of whom they wanted to send a message about. We limited the gossip opportunities to four possible targets as in each round, participants could observe the decisions of four other players. It was, however, not required to send gossip about these participants, since boxes could have been filled in with any ID. Participants were assisted in their gossip choices by the read-only display of their 'reputation table' on their screen. For each target, participants could select positive, neutral, or negative emoticons as the gossip message. Sending gossip was optional, and it was free of charge.

Gossip messages were not anonymous. On the fourth screen, incoming gossip messages became readable along with the ID of the sender. On the same screen, participants could assign or update reputation scores to all other (19) participants. More precisely, the instruction on the screen asked participants to privately evaluate how trustable they think others are on a scale of 0 to 100. Reputation scores were private assessments and participants were informed that the scores they gave to others were only visible to them. Previously given scores were displayed as reference values. On the fifth and last screen of each round, participants learned their own average score received from everyone else along with their payoff in the given round.

2.4.3. Design

The experiment in Phase II followed a 2x2 between-subject design. Four treatment conditions were constructed by the combination of two manipulations, both of which addressed the private reputational component of the experiment. First, we manipulated the scarcity of reputational rewards, second, we modified whether reputation had direct monetary effect on participants' payoffs. The scarcity of reputations was manipulated by the way participants could distribute reputation scores to others (on a scale between 0 and 100). Participants could either had a fixed budget of reputation scores (scarcity) or there was no ceiling on the distributable scores (abundance). In the abundance treatment, a participant could assign any number between 0 and 100 to each participant, a maximum of 1900 points was distributable (100N-1) in total. Theoretically, it can happen that everyone achieves a maximum reputation of 100. In the scarcity treatment, we limited the distributable scores to 950 (50N-1). If a subject here wanted to give 100 points to someone, then only 850 points have remained to be shared among the other 18 participants. Direct benefits for reputation was manipulated as reputation scores were either symbolic (not paid) or were incentivized financially (paid well). In the latter case, participants received the payoffs from the PD games and nothing more or less if they received 50 reputation points (the midscale value) from other participants on average. Otherwise, a one-unit decrease/increase from the default value of 50, reduced/increased their payments by HUF 20 (EUR 0.06). For instance, if all participants gave zero reputation to someone, then the receiver's payment was decreased by HUF 1000 (EUR 3.12). The four experimental groups are the combination of these two manipulations. In one of the condition reputations were not paid, and individuals could obtain reputation scores as many as they want out of the 100 (*abundance-not paid*). In the second case, accessible reputation was unlimited as in the previous condition, but payoffs were affected by players' average reputation scores (*abundance-paid well*). When unpaid reputation was scarce players reputation might have been undermined if others obtained more reputation than 50 (*scarcity-not paid*). Under the condition with limited access participants not just might ended up with bad reputation but they also paid fine because of it (*scarcity-paid well*).

2.5. Supplementary

A1.

Table S1. Categorization of potential explanations behind honest and dishonest gossip.

	Honest	Dishonest
Positive	Shared interest Strong reciprocity – reward for cooperators	Nepotism, strategic alliance, group-serving
Negative	Strong reciprocity – punishment of defectors	Conflict of interest – hurting rivals

A2.

Table S1 Estimated coefficients with multilevel mixed-effects logistic regressions (separate models for treatments)

Dependent variable: Prisoner's Dilemma choice **0 = defect, 1 = cooperate**	**ABUNDANCE – NOT PAID (A – NP)**	**ABUNDANCE –PAID WELL (A –PW)**	**SCARCITY – NOT PAID (S – NP)**	**SCARCITY-PAID WELL (S –PW)**
baseline: Round 1-5				
Round 6	0.1518	0.4268	0.4132	0.8665**
	[0.3552]	[0.3131]	[0.3570]	[0.3358]
Round 7-22	-0.3691*	-0.0514	-0.143	0.229
	[0.1622]	[0.1552]	[0.1758]	[0.1604]
constant	-1.6697***	-1.3349***	-2.3969***	-2.0165***
	[0.3070]	[0.3054]	[0.4578]	[0.3346]
lns1_1_1	0.4875**	0.5104***	0.8784***	0.5807***
	[0.1487]	[0.1444]	[0.1678]	[0.1542]
Nr. of groups	40	40	40	40
N of obs.	1700	1700	1700	1718

Note*: * p<.05; ** p<.01; *** p<.001, standard errors are in brackets, lns1_1_1: random intercept variance between subjects*

Table S2 Analysis of Variance of Gossip between treatments

Source	SS	df	MS	F	Prob > F
Between groups	86.28	3	28.76	36.88	0

Within groups	4403.92	5648	0.78
Total	4490.19	5651	0.79

Note: Bartlett's test for equal variances: $\chi 2(3)$ = 8.7147 Prob>$\chi 2$ = 0.033

Table S3 Comparison of Gossip by treatment (Bonferroni)

	ABUNDANCE - NOT PAID (A–NP)	ABUNDANCE - PAID WELL (A–PW)	SCARCITY - NOT PAID (S–NP)
ABUNDANCE - PAID WELL (A–PW)	0.32		
	0		
SCARCITY - NOT PAID (S–NP)	0.08	-0.24	
	0.09	0	
SCARCITY - PAID WELL (S–PW)	0.05	-0.27	-0.03
	0.846	0	1

Table S4 Summary of Gossip by treatment

	Mean	Std. Dev.
ABUNDANCE - NOT PAID (A–NP)	2.07	0.91
ABUNDANCE - PAID WELL (A–PW)	2.39	0.84
SCARCITY - NOT PAID (S–NP)	2.15	0.90
SCARCITY - PAID WELL (S–PW)	2.12	0.88
Total	2.19	0.89

Figure S1 Gossip values by PD outcomes and treatment

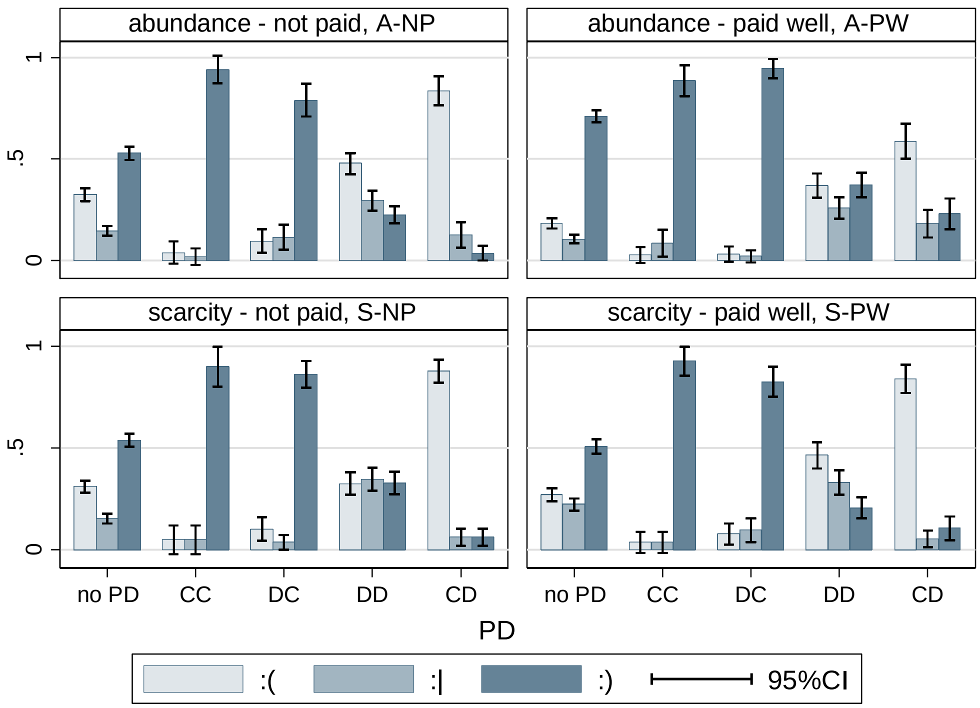


Table S5 Random-effects ordered logistic models

Dependent variable: Ego's gossip choice about Alter 1 = :(2 = :| 3 = :)	**Model 1**	**Model 2**	**Model 3**
baseline: ABUNDANCE - NOT PAID (A–NP)			
ABUNDANCE - PAID WELL (A–PW)	0.9570**	0.6639	0.7738
	[0.3615]	[0.3951]	[0.4720]
SCARCITY - NOT PAID (S–NP)	0.2038	0.4114	0.3218
	[0.3600]	[0.3923]	[0.4480]
SCARCITY - PAID WELL (S–PW)	0.2183	0.1435	0.472
	[0.3620]	[0.3950]	[0.4632]
PD decision (played)			
baseline: Not a PD Partner			
Ego cooperate – Alter cooperate		3.3595***	3.6162***
		[0.2915]	[0.6533]
Ego cooperate – Alter defect		-2.7489***	-3.2148***
		[0.1603]	[0.3288]
Ego defect – Alter cooperate		2.4422***	2.3056***
		[0.1790]	[0.3094]
Ego defect – Alter defect		-0.9184***	-0.9155***
		[0.0940]	[0.1756]
PD decision (observed)			
baseline: Not observed			
Alter cooperate		1.3421***	1.3448***
		[0.1540]	[0.2904]
Alter defect		-0.7453***	-0.8634***
		[0.0951]	[0.1851]
Reputation score distributed to Alter in the previous round		0.0208***	0.0241***
		[0.0014]	[0.0026]
Round		-0.0034	-0.0036
		[0.0075]	[0.0075]
PD decision (played) – treatment interaction			
ABUNDANCE - PAID WELL (A–PW) * Ego cooperate – Alter cooperate			-0.9999
			[0.7865]
SCARCITY - NOT PAID (S–NP) * Ego cooperate – Alter cooperate			0.9024

	[1.0648]
SCARCITY - PAID WELL (S–PW) * Ego cooperate – Alter cooperate	-0.0085
	[0.8805]
ABUNDANCE - PAID WELL (A–PW) * Ego cooperate – Alter defect	0.9974*
	[0.4312]
SCARCITY - NOT PAID (S–NP) * Ego cooperate – Alter defect	0.0052
	[0.4858]
SCARCITY - PAID WELL (S–PW) * Ego cooperate – Alter defect	0.5758
	[0.4740]
ABUNDANCE - PAID WELL (A–PW) * Ego defect – Alter cooperate	0.7942
	[0.6395]
SCARCITY - NOT PAID (S–NP) * Ego defect – Alter cooperate	0.0426
	[0.4643]
SCARCITY - PAID WELL (S–PW) * Ego defect – Alter cooperate	0.1958
	[0.4604]
ABUNDANCE - PAID WELL (A–PW) * Ego defect – Alter defect	-0.2685
	[0.2650]
SCARCITY - NOT PAID (S–NP) * Ego defect – Alter defect	0.2223
	[0.2514]
SCARCITY - PAID WELL (S–PW) * Ego defect – Alter defect	-0.0476
	[0.2660]
PD decision (observed) – treatment interaction	
ABUNDANCE - PAID WELL (A–PW) * Alter cooperate	-0.8523*
	[0.4217]
SCARCITY - NOT PAID (S–NP) * Alter cooperate	-0.027
	[0.4470]
SCARCITY - PAID WELL (S–PW) * Alter cooperate	0.636
	[0.4190]
ABUNDANCE - PAID WELL (A–PW) * Alter defect	0.3106
	[0.2803]
SCARCITY - NOT PAID (S–NP) * Alter defect	0.2074

			[0.2574]
SCARCITY - PAID WELL (S–PW) * Alter defect			0.0194
			[0.2684]
Reputation score distributed to Alter in the previous round – treatment interaction			
ABUNDANCE - PAID WELL (A–PW) * Reputation score distributed to Alter in the previous round			-0.0033
			[0.0037]
SCARCITY - NOT PAID (S–NP) * Reputation score distributed to Alter in the previous round			-0.0002
			[0.0038]
SCARCITY - PAID WELL (S–PW) * Reputation score distributed to Alter in the previous round			-0.0095*
			[0.0038]
cut1	-0.7312**	-0.3212	-0.2387
	[0.2529]	[0.3048]	[0.3368]
cut2	0.2786	1.0574***	1.1494***
	[0.2528]	[0.3052]	[0.3372]
sigma2_u	2.3562***	2.7842***	2.8148***
	[0.3200]	[0.3816]	[0.3867]
Nr. of groups	156	156	156
Nr. of obs.	5599	5480	5480

Note: * p<.05; ** p<.01; *** p<.001, standard errors are in brackets, sigma2_u: variance component attributable to subjects

Table S6 Analysis of Variance of Reputation Score between treatments

Source	SS	df	MS	F	Prob > F
Between groups	7275618.90	3	2425206.30	3110.80	0
Within groups	40372794.30	51786	779.61		
Total	47648413.20	51789	920.05		

Note: Bartlett's test for equal variances: chi2(3) = 1.8e+03 Prob>chi2 = 0.000

Table S7 Comparison of Reputation Score by treatment (Bonferroni)

	ABUNDANCE - NOT PAID (A–NP)	ABUNDANCE - PAID WELL (A–PW)	SCARCITY - NOT PAID (S–NP)
ABUNDANCE - PAID WELL (A–PW)	22.68		
	0		
SCARCITY - NOT PAID (S–NP)	-8.56	-31.24	
	0	0	
SCARCITY - PAID WELL (S–PW)	-2.68	-25.36	5.88
	0	0	0

Table S8 Summary of Reputation Score by treatment

	Mean	Std. Dev.
ABUNDANCE - NOT PAID (A–NP)	45.68	30.52
ABUNDANCE - PAID WELL (A–PW)	68.36	31.97
SCARCITY - NOT PAID (S–NP)	37.12	24.76
SCARCITY - PAID WELL (S–PW)	43.00	23.49
Total	48.53	30.33

Table S9 Multilevel mixed-effects linear regression

Dependent variable: Reputation score distributed to Alter between 0 – 100	**Model 1**	**Model 2**	**Model 3**	**Model 4**
baseline: ABUNDANCE - NOT PAID (A–NP)				
ABUNDANCE - PAID WELL (A–PW)	22.6828***	6.7665***	6.9360***	6.9257***
	[3.5550]	[1.1243]	[1.1258]	[1.1265]
SCARCITY - NOT PAID (S–NP)	-8.5606*	-2.6015*	-2.7537*	-2.8260*
	[3.5550]	[1.1227]	[1.1243]	[1.1249]
SCARCITY - PAID WELL (S–PW)	-2.6835	-0.9127	-1.2385	-1.3329
	[3.5550]	[1.1225]	[1.1238]	[1.1244]
baseline: Not a PD Partner				
Ego cooperate – Alter cooperate		16.0085***	16.0042***	15.9881***
		[0.8419]	[0.8410]	[0.8409]
Ego cooperate – Alter defect		-12.3325***	-12.3404***	-12.3412***
		[0.5039]	[0.5033]	[0.5033]
Ego defect – Alter cooperate		8.7767***	8.8255***	8.8222***
		[0.5353]	[0.5347]	[0.5347]
Ego defect – Alter defect		-2.9128***	-2.8934***	-2.8946***
baseline: Not observed				
Alter cooperate		4.2662***	4.3118***	4.3009***
		[0.4490]	[0.4486]	[0.4485]
Alter defect		-1.3760***	-1.3698***	-1.3734***
		[0.2556]	[0.2554]	[0.2554]
Reputation score distributed to Alter in the previous round		0.7325***	0.7327***	0.7328***
		[0.0030]	[0.0030]	[0.0030]
Round		0.0061	0.0069	0.0066
		[0.0141]	[0.0140]	[0.0140]
baseline: No gossip				
:(		-6.9627***	-7.0779***	-7.1236***
		[0.3776]	[0.6896]	[0.6900]
:|		-1.7248***	-0.854	-0.9002
		[0.5144]	[1.0018]	[1.0021]
:)		7.1155***	5.3406***	5.2942***
		[0.3036]	[0.6318]	[0.6323]
Nr. of gossip sent by Alter		0.3140**	0.3117**	-0.0409

		[0.1171]	[0.1170]	[0.2320]
Gossip – treatment interaction				
:(* ABUNDANCE - PAID WELL (A–PW)			-5.5211***	-5.5107***
			[1.1143]	[1.1150]
:(* SCARCITY - NOT PAID (S–NP)			2.4907*	2.5652**
			[0.9843]	[0.9851]
:(* SCARCITY - PAID WELL (S–PW)			1.8458	1.9457
			[1.0472]	[1.0480]
:| * ABUNDANCE - PAID WELL (A–PW)			-5.5822***	-5.5717***
			[1.5266]	[1.5271]
:| * SCARCITY - NOT PAID (S–NP)			0.7912	0.8671
			[1.4066]	[1.4071]
:| * SCARCITY - PAID WELL (S–PW)			0.1373	0.2359
			[1.4096]	[1.4101]
:) * ABUNDANCE - PAID WELL (A–PW)			1.1936	1.2037
			[0.8351]	[0.8362]
:) * SCARCITY - NOT PAID (S–NP)			0.8872	0.9636
			[0.8643]	[0.8653]
:) * SCARCITY - PAID WELL (S–PW)			5.7895***	5.8885***
			[0.9253]	[0.9261]
Nr. of gossip sent by Alter – treatment interaction				
ABUNDANCE - PAID WELL (A–PW)				0.0692
				[0.3290]
SCARCITY - NOT PAID (S–NP)				0.563
				[0.3199]
SCARCITY - PAID WELL (S–PW)				0.8291*
				[0.3436]
constant	45.6831***	12.0364***	12.0885***	12.1354***
	[2.5138]	[0.8313]	[0.8320]	[0.8323]
lns1_1_1	2.7630***	1.5981***	1.5976***	1.5975***
	[0.0563]	[0.0584]	[0.0584]	[0.0584]
lnsig_e	3.1350***	2.7503***	2.7491***	2.7490***
	[0.0031]	[0.0031]	[0.0031]	[0.0031]
Nr. of groups	160	160	160	160
Nr. of obs.	51790	51790	51790	51790

Note: * p<.05; ** p<.01; *** p<.001, standard errors are in brackets, lns1_1_1: random intercept variance between subjects, lnsig_e: random intercept variance within subjects

Table S10 Multilevel mixed-effects linear regression

Dependent variable: Reputation score distributed to Alter between 0 – 100	ABUNDANCE – NOT PAID (A – NP)	ABUNDANCE –PAID WELL (A –PW)	SCARCITY – NOT PAID (S – NP)	SCARCITY- PAID WELL (S –PW)
baseline: Not a PD Partner				
Ego cooperate – Alter cooperate	20.7174***	10.9406*	32.2006***	12.9391**
	[5.7471]	[5.1270]	[6.0023]	[4.7300]
Ego cooperate – Alter defect	-12.9081***	-16.5840***	-17.5938***	-7.2245*
	[3.2688]	[3.2400]	[3.0267]	[3.3525]
Ego defect – Alter cooperate	10.8297***	11.7203***	9.3753**	6.0602
	[3.0598]	[3.3729]	[3.2897]	[4.0470]
Ego defect – Alter defect	-5.0406**	-6.2372**	-0.6637	-3.4172
	[1.7044]	[2.2551]	[1.6485]	[1.8787]
baseline: Not observed				
Alter cooperate	4.3346	3.4113	7.7952**	0.8863
	[2.4549]	[2.6545]	[2.5551]	[2.7162]
Alter defect	-0.2387	-0.3301	-2.5172	-2.3225
	[1.5388]	[1.7692]	[1.4013]	[1.7585]
Reputation score distributed to Alter in the previous round	0.7125***	0.5545***	0.6068***	0.5297***
	[0.0194]	[0.0237]	[0.0214]	[0.0246]
Round	-0.1533	0.0149	0.2071**	-0.0517
	[0.0840]	[0.1002]	[0.0800]	[0.0960]
Gossip Partner’s reputation	0.1131***	0.1660***	0.1770***	0.2091***
	[0.0301]	[0.0353]	[0.0350]	[0.0410]
baseline: No gossip				
:(	-3.2991	1.8383	1.4254	3.0458
	[2.2328]	[3.4742]	[2.1424]	[2.7923]
:|	1.0749	3.5632	6.4906*	-1.3369
	[2.5620]	[4.7966]	[2.5268]	[3.3972]
:)	10.9064***	17.6405***	12.7418***	24.1751***
	[2.1246]	[2.9063]	[2.1221]	[2.5961]
Gossip Partner’s reputation - Gossip interaction				
:(	-0.1006**	-0.1771***	-0.1258**	-0.1863***
	[0.0358]	[0.0451]	[0.0433]	[0.0556]
:|	-0.0588	-0.0946	-0.1533**	0.0246
	[0.0457]	[0.0608]	[0.0545]	[0.0682]
:)	-0.1394***	-0.1102**	-0.1326**	-0.2813***

	[0.0350]	[0.0369]	[0.0418]	[0.0511]
Nr. of gossip sent by Alter	-0.3312	1.1093*	0.8784*	1.0476*
	[0.4400]	[0.5160]	[0.4228]	[0.4671]
_cons	10.8726***	16.3587***	3.2347	10.7976***
	[2.2455]	[2.8709]	[2.0194]	[2.3715]
lns1_1_1	1.7561***	1.9212***	1.3250***	1.3582***
	[0.1480]	[0.1448]	[0.1805]	[0.1720]
lnsig_e	2.8789***	3.0607***	2.8745***	2.9488***
	[0.0155]	[0.0157]	[0.0152]	[0.0164]
Nr. of groups	160	160	160	160
Nr. of obs.	2123	2087	2213	1910

Note: * p<.05; ** p<.01; *** p<.001, standard errors are in brackets, Nr. of obs.: Table contains observations where Ego received gossip, lns1_1_1: random intercept variance between subjects, lnsig_e: random intercept variance within subjects

Figure S2. The usability of the reputation systems by treatments

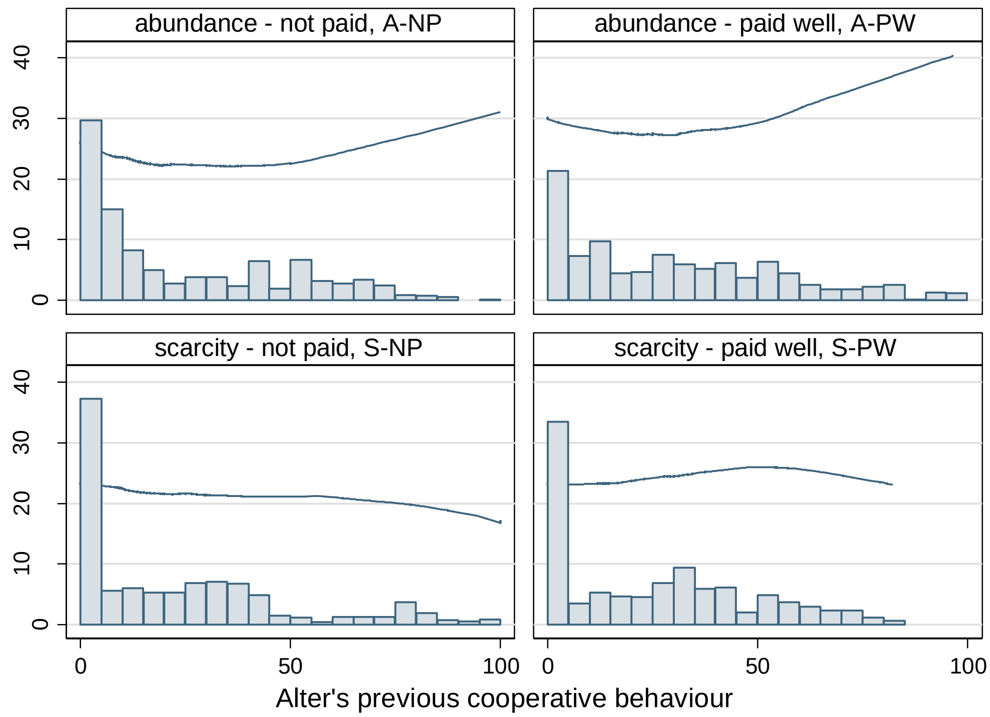


Figure A2S3 depicts the average degree of cooperation with alter as a function of alter's average cooperation level and the distribution of cooperative behaviour during the experiment. ABUNDANCE - PAID WELL (A–PW) has the lowest number of continuous defectors and SCARCITY - NOT PAID (S–NP) has the highest. There were no continuous co-operators in SCARCITY - PAID WELL (S–PW). The line graph shows ego's decision as a function of alter's behaviour until that round. The highest fitted slope (using Locally Weighted Scatterplot Smoothing, LOWESS method) between alter's previous behaviour and the decision against alter was found in SCARCITY - PAID WELL (S–PW) and ABUNDANCE - NOT PAID (A–NP), so participants made the most accurate decisions in these treatments.

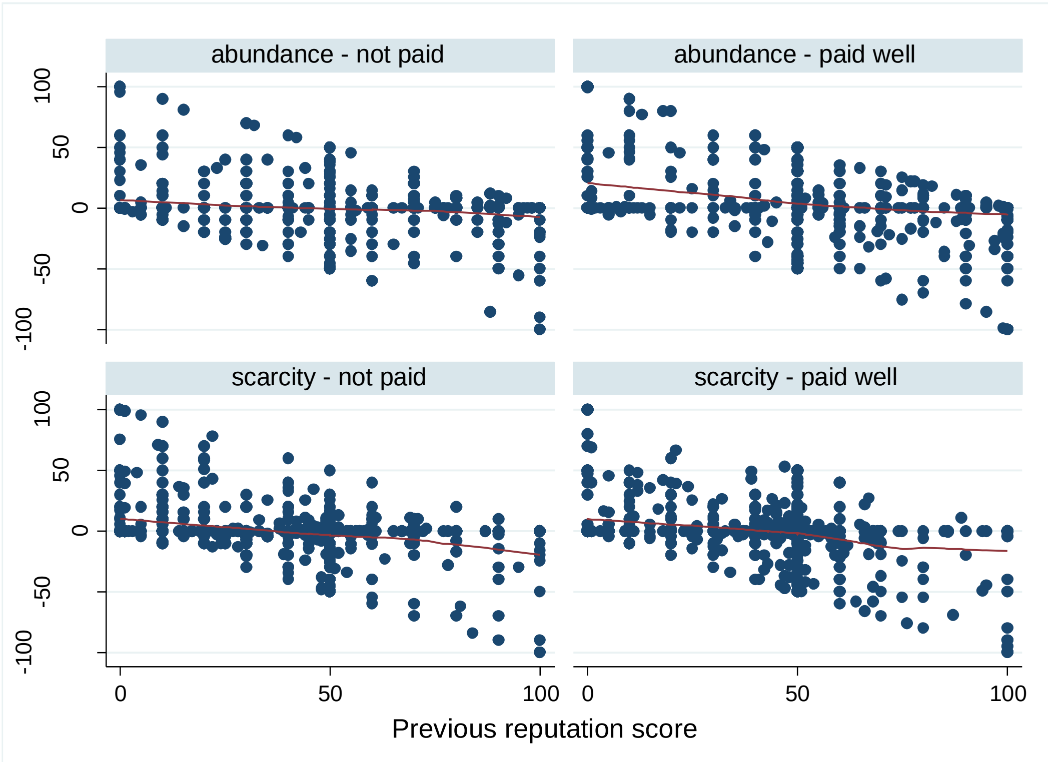
Table S11 Multilevel mixed-effects linear regression

	Model 1	Model 2
Dependent variable:	Reputation distributed to Alter between 0 – 100	Changes in Alter's reputation
Treatment (baseline: ABUNDANCE - NOT PAID (A–NP))		
ABUNDANCE - PAID WELL (A–PW)	13.2715***	13.3583***
	[3.2063]	[3.2132]
SCARCITY - NOT PAID (S–NP)	0.2067	0.327
	[2.6061]	[2.6315]
SCARCITY - PAID WELL (S–PW)	3.1259	2.9222
	[2.8193]	[2.8334]
Focal player's reputation	0.8128***	-0.1844***
	[0.0382]	[0.0383]
Focal player's reputation – treatment interaction		
Focal player's reputation * ABUNDANCE - PAID WELL (A–PW)	-0.1715**	-0.1707**
	[0.0560]	[0.0561]
Focal player's reputation * SCARCITY - NOT PAID (S–NP)	-0.0292	-0.0335
	[0.0553]	[0.0557]
Focal player's reputation * SCARCITY - PAID WELL (S–PW)	-0.1536**	-0.1576**
	[0.0563]	[0.0571]
Opponent's reputation	-0.0025	0.0005
	[0.0383]	[0.0384]
Opponent's reputation – treatment interaction		
Opponent's reputation * ABUNDANCE - PAID WELL (A–PW)	0.0558	0.0524
	[0.0562]	[0.0563]
Opponent's reputation * SCARCITY - NOT PAID (S–NP)	0.0566	0.0547
	[0.0523]	[0.0527]
Opponent's reputation * SCARCITY - PAID WELL (S–PW)	-0.0146	-0.0116
	[0.0558]	[0.0559]
Focal player's reputation – Opponent's reputation interaction	0.0001	0.0524
	[0.0006]	[0.0563]
Focal player's reputation – Opponent's reputation – treatment interaction		
Focal player's reputation – Opponent's reputation * ABUNDANCE - PAID WELL (A–PW)	-0.0001	0.0000
	[0.0008]	[0.0008]
Focal player's reputation – Opponent's reputation * SCARCITY - NOT PAID (S–NP)	-0.0024*	-0.0023*
	[0.0010]	[0.0010]
Focal player's reputation – Opponent's reputation * SCARCITY - PAID WELL (S–PW)	0.0011	0.0013
	[0.0011]	[0.0011]
Focal player's decision (baseline: D)		
C	12.2038***	12.1294***
	[3.6599]	[3.6601]
Focal player's decision * treatment interaction		
C * ABUNDANCE - PAID WELL (A–PW)	0.2514	0.8586
	[5.8325]	[5.8384]
C * SCARCITY - NOT PAID (S–NP)	3.9982	4.1214
	[4.7190]	[4.7431]
C * SCARCITY - PAID WELL (S–PW)	-14.2599**	-14.3638**
	[5.2902]	[5.3146]
Focal player's decision * Focal player's reputation interaction		
C * Focal player's reputation	-0.1277	-0.1306
	[0.0790]	[0.0790]

Focal player's decision - Focal player's reputation - treatment interaction		
C * Focal player's reputation * ABUNDANCE - PAID WELL (A–PW)	0.033	0.0231
	[0.1148]	[0.1150]
C * Focal player's reputation * SCARCITY - NOT PAID (S–NP)	0.0262	0.027
	[0.1025]	[0.1027]
C * Focal player's reputation * SCARCITY - PAID WELL (S–PW)	0.3065**	0.3064**
	[0.1101]	[0.1107]
Focal player's decision – Opponent's reputation interaction		
C * Opponent's reputation	-0.0061	0.0025
	[0.0793]	[0.0800]
Focal player's decision – Opponent's reputation – treatment interaction		
C * Opponent's reputation * ABUNDANCE - PAID WELL (A–PW)	-0.1409	-0.1592
	[0.1148]	[0.1157]
C * Opponent's reputation * SCARCITY - NOT PAID (S–NP)	-0.0512	-0.0541
	[0.1099]	[0.1107]
C * Opponent's reputation * SCARCITY - PAID WELL (S–PW)	0.3804***	0.3619**
	[0.1148]	[0.1155]
Focal player's decision – Focal player's reputation – Opponent's reputation		
C * Opponent's reputation * Focal player's reputation	0.0000	-0.0001
	[0.0012]	[0.0012]
Focal player's decision – Focal player's reputation – Opponent's reputation – treatment interaction		
C * Opponent's reputation * Focal player's reputation * ABUNDANCE - PAID WELL (A–PW)	0.0015	0.0018
	[0.0016]	[0.0016]
C * Opponent's reputation * Focal player's reputation * SCARCITY - NOT PAID (S–NP)	-0.0004	-0.0004
	[0.0019]	[0.0019]
C * Opponent's reputation * Focal player's reputation * SCARCITY - PAID WELL (S–PW)	-0.0073***	-0.0070***
	[0.0020]	[0.0020]
_cons	7.4015***	7.3181***
	[1.9088]	[1.9195]
lns1_1_1	1.6794***	1.6793***
	[0.0857]	[0.0862]
lnsig_e	2.9001***	2.8986***
	[0.0099]	[0.0100]
Nr. of groups	160	160
Nr. of obs.	5347	5254

Note: * p<.05; ** p<.01; *** p<.001, standard errors are in brackets, Nr. of obs.: Table contains observations where Ego observes a PD game as a third party, lns1_1_1: random intercept variance between subjects, lnsig_e: random intercept variance within subjects

Figure S3. Previous reputational scores and reputation updates by treatments



Notes: The Figure shows how previous reputational scores affect reputation updates by treatments. The red line is fitted by using a locally weighted regression (lowess) of changes in scores on reputation scores in the previous round.

Table S12 Overall correlation between behaviour, gossip, reputation score by treatment

		Alter's reputation score		Degree of cooperation with Alter		Gossip about Alter	
Control (Round 1-5)	Alter's cooperation			0.04	0.65		
	Alter's reputation score						
	Degree of cooperation with Alter						
A–NP	Alter's cooperation	0.57	0.00	0.21*	0.19	0.71	0.00
	Alter's reputation score			0.55	0.00	0.74	0.00
	Degree of cooperation with Alter					0.38	0.01
A–PW	Alter's cooperation	0.75	0.00	0.16	0.31	0.63	0.00
	Alter's reputation score			-0.01	0.93	0.64	0.00
	Degree of cooperation with Alter					0.25	0.13
S–NP	Alter's cooperation	0.43	0.01	0.04	0.83	0.71	0.00
	Alter's reputation score			-0.29	0.07	0.64	0.00
	Degree of cooperation with Alter					-0.10	0.55
S–PW	Alter's cooperation	0.69	0.00	0.38	0.02	0.69	0.00
	Alter's reputation score			0.29	0.07	0.82	0.00
	Degree of cooperation with Alter					0.26	0.10

Note: *significant from Round 10 (corr=0.34, sig=0.03)

Figure S4 Graphical illustration of correlations between behaviour, gossip, reputation score by treatment

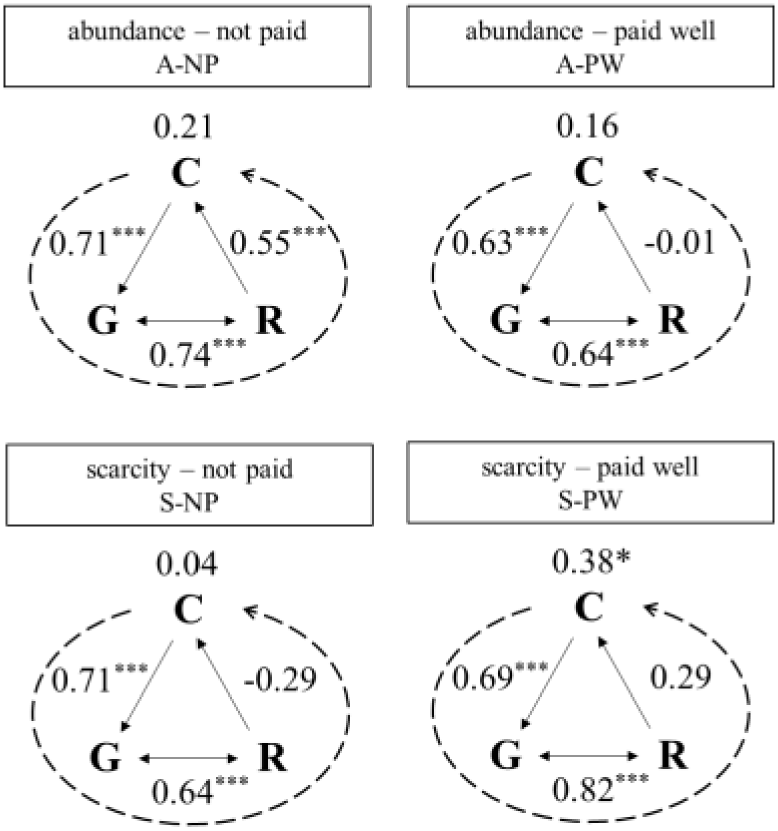


Table S13 Multilevel mixed-effects logistic regression

Dependent variable: Prisoner's Dilemma choice 0 = defect, 1 = cooperate	ABUNDANCE - NOT PAID (A–NP) Model 1	ABUNDANCE - PAID WELL (A–PW) Model 1	SCARCITY - NOT PAID (S–NP) Model 1	SCARCITY - PAID WELL (S–PW) Model 1	Joint model
Alter's previous cooperative behaviour	1.1662***	0.3444	0.352	0.9393**	
	[0.3392]	[0.2999]	[0.3512]	[0.3426]	
Alter's reputation score					0.0382***
					[0.0041]
baseline: ABUNDANCE - NOT PAID (A–NP)					
ABUNDANCE - PAID WELL (A–PW)					3.6389**
					[1.1171]
SCARCITY - NOT PAID (S–NP)					-0.0599
					[0.9809]
SCARCITY - PAID WELL (S–PW)					1.5402
					[1.1582]
Alter's Reputation score – treatment interaction					
ABUNDANCE - PAID WELL (A–PW) * Alter's Reputation score					-0.0278***
					[0.0051]
SCARCITY - NOT PAID (S–NP) * Alter's Reputation score					-0.0088
					[0.0060]
SCARCITY - PAID WELL (S–PW) * Alter's Reputation score					0.0021
					[0.0061]
Ego's Reputation score					-0.0155
					[0.0141]
Ego's Reputation score – treatment interaction					
ABUNDANCE - PAID WELL (A–PW) * Ego's Reputation score					-0.0223
					[0.0184]
SCARCITY - NOT PAID (S–NP) * Ego's Reputation score					0.0178
					[0.0199]
SCARCITY - PAID WELL (S–PW) *					-0.02

Ego's Reputation score					
					[0.0237]
Round					-0.0314***
					[0.0094]
Ego's cooperation level in Round 1-5					4.6916***
					[0.6140]
_cons	-2.6362***	-1.5419***	-2.8155***	-1.9954***	-4.4143***
	[0.3855]	[0.3179]	[0.4941]	[0.3176]	[0.7913]
lns1_1_1	0.7011***	0.5803***	0.9446***	0.5401***	0.6297***
	[0.1716]	[0.1508]	[0.1729]	[0.1576]	[0.0822]
Nr. of groups	40	40	40	40	120
N	1274	1331	1257	1228	5090

Note: * p<.05; ** p<.01; *** p<.001, standard errors are in brackets, lns1_1_1: random intercept variance between subjects

A3. Detailed discussion of treatment effects

In the setting with scarce reputational resources that could directly be translated to monetary gains (S-WP), we expected an intense competition between participants by either increasing their own cooperative behaviour or by wrecking the position of others. Therefore, we expected that cooperation will be higher, and competitors will use the opportunity to worsen the position of others by dishonest gossip. Although only in the short term, it seems that in this treatment, participants took this competition more seriously and strived more for achieving their own reputations by cooperation. In addition, we observed less positive gossiping about individuals with high reputation. The dissemination of false information not just contributes to the deterioration of reputation, but it hinders the reliability of the reputation system. Since individuals are only willing to take the risk of cooperative behaviour if the reputation system is reliable, this can have a negative effect on cooperation. Although strategic cooperation disappeared in the long run, the collapse of the reputation system did not happen in the SCARCITY – PAID WELL S-PW treatment. Reputations have preserved their credibility over time despite the possibility of misinformation maybe because positive gossip was more credible and have been taken into greater account in scoring.

The reputation system was unable to increase cooperative behaviour where limited reputational resources were available, but there was no external motivation for reputation (SCARCITY - NOT PAID, S–NP). Since punishment was more lenient and symbolic in a way that it has not been accompanied by payoff reduction, competition may have not been taken seriously by participants. We also found a slightly different pattern in participants' reliance on gossiping as negative gossip resulted in less score reduction than in other treatments. This may follow from the fact that improvement of relative position in this treatment can be reached not just by cooperation but by sending negative gossip about others, therefore individuals did not entirely believe them.

The use of external incentives without competition (ABUNDANCE - PAID WELL, A–PW) could have hampered the development of a trusted reputation system in various ways. Given unlimited reputational resources, individuals tried to encourage cooperation by sending positive messages and giving high reputation points for everyone, making it impossible to use the reputational system to differentiate cooperative intentions of others. Although participants tended to balance this positivity by reducing reputation scores in a greater extent of those who were a target of a neutral or a negative gossip, they were less dependent on reputation scores during their PD decisions.

Although we do not observe strategic reputation building, the reputational system could supervise decisions where reputation building was not incentivized externally, and the achievement of good reputation was available for everyone (ABUNDANCE - NOT PAID, A–NP), as we found an association between players' decision in the PD game and partners' previous behaviour.

A4. Instructions of the experiment

Welcome to the decision-making experiments organized by the Corvinus University of Budapest!

The decision-making experiments carried out by the Hungarian Academy of Sciences, Centre for Social Sciences, "Lendület" Research Center for Educational and Network Studies (RECENS), led by Károly Takács, and supported by the European Research Council (ERC CoG 648693).

Please turn off your phone or completely turn it down!

In the following, instructions will appear on your screen about the experiment. You participate in the experiment together with people in this room. Instructions are the same for all participants. You get the most important instructions on paper as well. You can use them at any time during the experiment. Please do not take these instructions with you after the experiment, leave them on the table.

VERY IMPORTANT rule is that it is STRICTLY FORBIDDEN to talk to or to signal to others! Violation of this rule may result in disqualification from this experiment.

The experiment takes about 75 minutes. Your payoffs in the experiment will be paid at the end of the experiment.

The amount of your payoff depends on your own choices and the decisions of others. The precise calculation of your payment will be described later in more detail.

By pressing the \ "Next \" button, you agree that your answers will be exclusively and anonymously used for scientific research purposes.

Thank you in advance for your participation!

If you're ready, press the \ "Next \" button! Have fun and good luck!

INSTRUCTIONS

The experiment will consist of several rounds of decision-making. In each round, you will be paired with two other participants. Participants of the experiment will be identified with numbers ranging from 1 to 20.

It is important that you do not play with the same participants in each round! Both of your pair determines by a random number generator. The ID of your current pair will be displayed on your screen. The IDs will be kept confidential, neither during the experiment nor at the end of it will we not reveal which identifier participants belonged to.

Each round is important for your final payoff!

1 round in the first phase and 5 rounds in the second phase will be selected using a random number generator. The average winnings in these rounds will be your final payoff. We add everyone 1000 forints as a bonus. Payoffs will be rounded up to HUF 100.

Each pair faces with the following decision-making options. Two options are provided: the two options are indicated with L and R. The amount that you win this round does not depend only on your decision, but also on your partner's decision. It is detailed in the following what payoff can be expected:

If both of you choose L: HUF 1500

If you choose L and your partner select R: HUF 2500

If you choose R and your partner select L: HUF 0

If both of you choose R: HUF 500

If you run out of time: HUF 0

It is important that all the information that you receive is real. The time available to you will be projected in the upper right corner of your screen.

When you are ready, please click on the "Next" button.

FURTHER INSTRUCTIONS

The same decision will be taken in the next rounds and the amounts that you can win are the same as previously.

In each round, you will randomly be paired with two other participants. It is therefore important that you do not play with the same participants in each round, but your pair is determined by a random number generator!

The change is that now the participants can be scored from 0 to 100 by you on the basis how reliable they are according to you A maximum of 950 points can be distributed. By default, the starting score is set to 50, as a neutral medium. (So long as you distribute more than the maximum points, you get an error message. If this will not be corrected in time, the total score will be rounded down proportionately.)

The points you receive will be taking into account in the final payoff!

The payoff for that run will be adjusted by your average point.

If the point you received on average corresponding to the neutral value of 50, then your payoff will be unchanged in the current round. In comparison, a one-unit decrease/increase in your average point reduces/increases your payments by HUF 20. For instance, if all other participants give you 0 point, then your payment decreases by HUF 1000. If all other participants give you 100 point, your payment increases by HUF 1000.

In addition to the scoring it is also a new element that one pair from the previous rounds will be randomly selected, and you get acquainted with the decisions they have made in that round.

Then, you'll be able to send a message to another randomly selected participant. The ID of the participant – to whom you can send the message – will appear on your screen. Then, optionally, you can select four participants, the four whom the message is about. There is no cost of sending a message.

The time available for your choice will display in the top right corner of the screen. Attention please! If you run out of time, it is considered that you did not want to send a message!

After all messages have been sent, messages you received will appear on your screen.

Now you can also modify the scores on the 100 points scale that you have given to others indicating how reliable they are.

When you are ready, please click on the "Next" button. Have fun and good luck!

A5. Screens of one round in the first phase (Round 1-5) of the experiment (with English translation)

Round 2	Remaining time (sec): 19

This round may count towards your final payment. Remember, your payoff is:
If both of you choose L: 1500
If you choose L and your partner chooses R: 0
If you choose R and your partner chooses L: 2500
If both of you choose R: 500

Your ID: 5

Your first Partner's ID: 2
Which option do you choose against her/him? L/R

Your second Partner's ID: 8
Which option do you choose against her/him? L/R

To confirm your decision, you need to click on the 'confirm' button.

Confirm

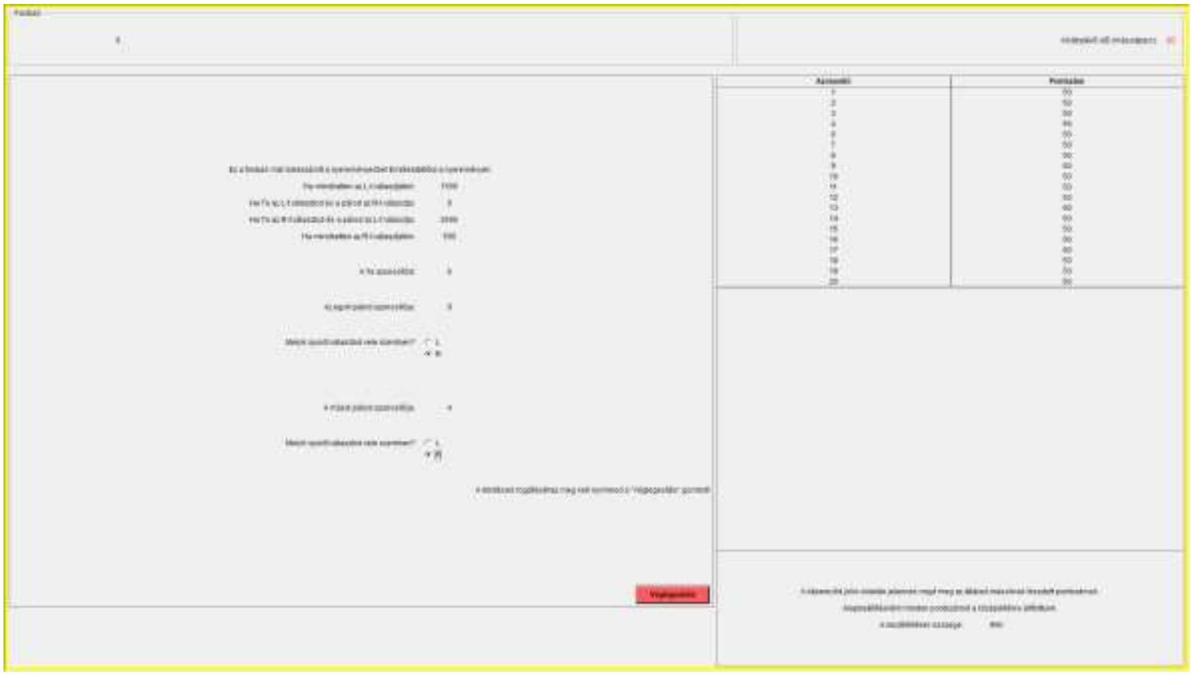
Your decision against your first partner in this round: R
As a reminder, your partner's ID: 2
Your partner's decision in this round: L
Your payoff in this round: 2500

Your decision against your second partner in this round: R
As a reminder, your partner's ID: 8
Your partner's decision in this round: R
Your payoff in this round: 500

Your combined payoff in this round (in HUF): 1500

OK

A6. Screens of one round in the second phase (Round 6-22) of the experiment (with English translation)



Round 6

Remaining time (sec): 20

This round may count towards your final payment. Remember, your payoff is:
If both of you choose L: 1500
If you choose L and your partner chooses R: 0
If you choose R and your partner chooses L: 2500
If both of you choose R: 500

Your ID: 5

Your first Partner's ID: 9
Which option do you choose against her/him? L/R

Your second Partner's ID: 4
Which option do you choose against her/him? L/R

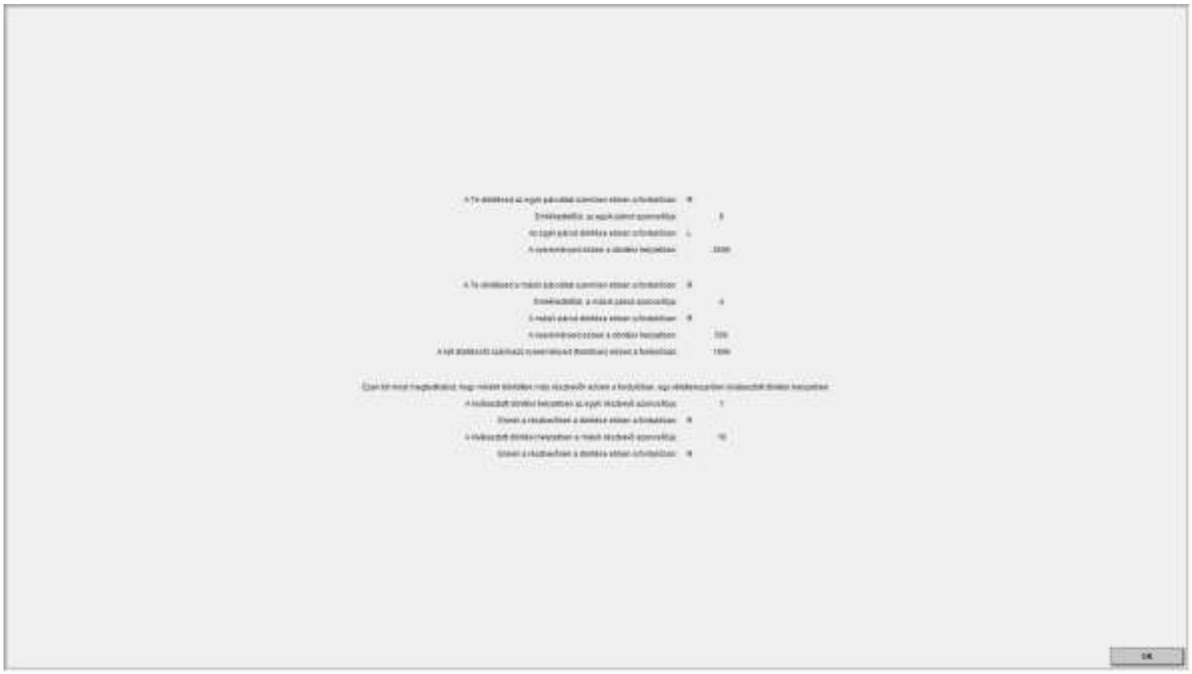
To confirm your decision, you need to click on the 'confirm' button.

Confirm

ID	Scores
1	50
2	50
3	50
4	50
6	50
7	50
8	50
9	50
10	50
11	50
12	50
13	50
14	50
15	50
16	50
17	50
18	50
19	50
20	50

The scores you give to others will appear on the right side of your screen.
At the beginning, all scores were set to mean.
The total default value is 950.*

Note: *From Round 7 the text has changed: 'The total score allocated by you is: … ' which text was missing in the 'abundance treatment'



Your decision against your first partner in this round: R
As a reminder, your partner's ID: 9
Your partner's decision in this round: L
Your payoff in this round: 2500

Your decision against your second partner in this round: R
As a reminder, your partner's ID: 4
Your partner's decision in this round: R
Your payoff in this round: 500

Your combined payoff in this round (in HUF): 1500

Moreover, now you can learn how other participants have decided in this round, in a randomly selected play:
The ID of one participant in the selected play: 1
The decision of this participant in this round: R
The ID of the other participant in the selected play: 10
The decision of this participant in this round: R

OK

Round 6	Remaining time (sec): 20

As a reminder, your ID is: 5
Now you can send messages to a randomly selected participant.
If you run out of time, it is considered that you did not want to send a message.

TO WHOM you send messages: 6

Now you can choose ABOUT WHO you want to send a message.
You can select a max. of 4 participants.
You cannot choose yourself and who you are sending the message to. You can select one participant only once.

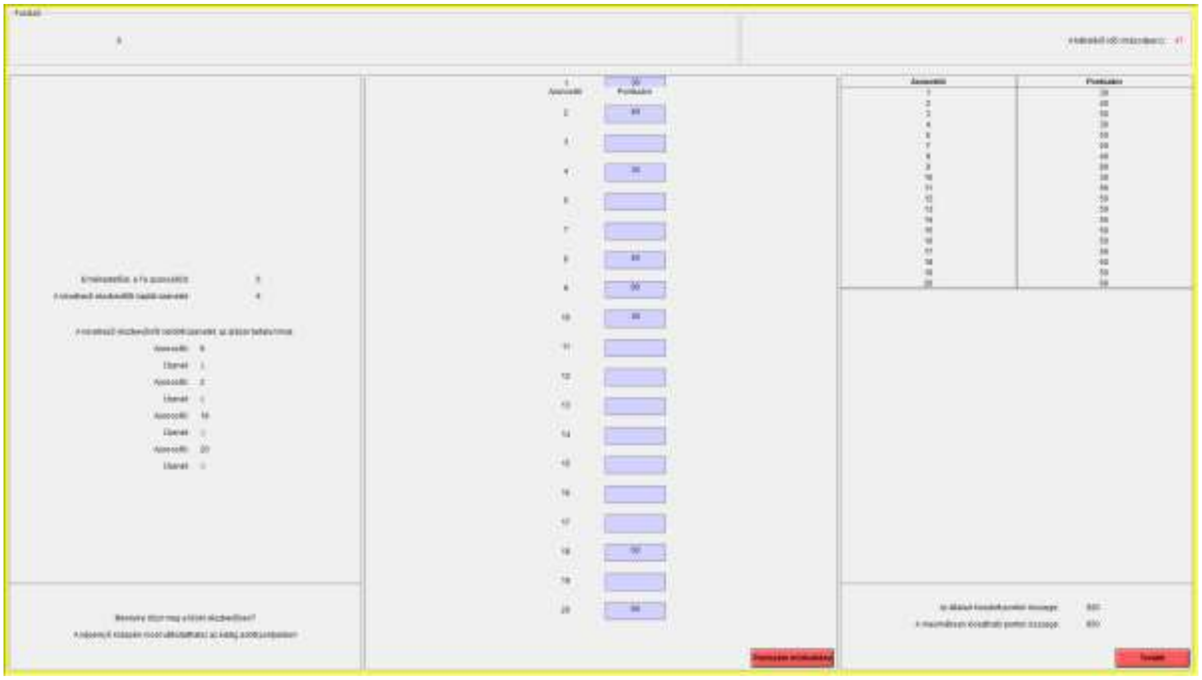
ABOUT WHO do you want to send a message? Write his/her ID here:
What message would you like to send about him/her? :) / :| / :(
ABOUT WHO do you want to send a message? Write his/her ID here:
What message would you like to send about him/her? :) / :| / :(
ABOUT WHO do you want to send a message? Write his/her ID here:
What message would you like to send about him/her? :) / :| / :(
ABOUT WHO do you want to send a message? Write his/her ID here:
What message would you like to send about him/her? :) / :| / :(

Next

ID	Scores
1	50
2	50
3	50
4	50
6	50
7	50
8	50
9	50
10	50
11	50
12	50
13	50
14	50
15	50
16	50
17	50
18	50
19	50
20	50

THESE SCORES ARE DEFAULT VALUES.
The total default value is 950.*

Note: *From Round 7 the text has changed: 'The total score allocated by you is: … The maximum scores you can allocate is 950.' The second sentence appeared only where scarcity was introduced.

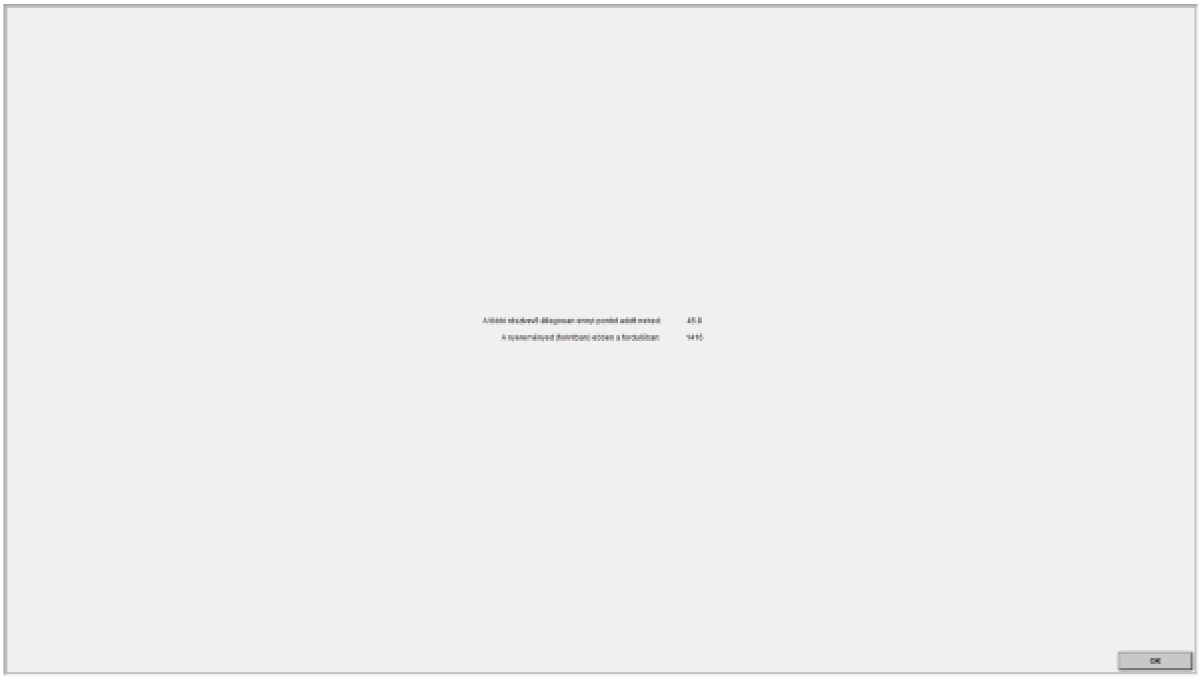


Round 6		Remaining time (sec): 20
As a reminder, your ID is: 5 You received messages FROM: 6 You received the following messages: ID: 8 Message: :(ID: 2 Message: :(ID: 18 Message: :) ID: 20 Message: :)	(see middle table below)	(see right table below)
How much do you trust other participants? Now, in the middle of the screen you can change the scores of other participants.		THESE SCORES ARE DEFAULT VALUES. The total default value is 950.

ID	Scores
1	30
2	40
3	_
4	30
6	–
7	_
8	40
9	80
10	_30
11	–
12	–
13	–
14	–
15	–
16	–
17	_
18	60
19	_
20	60

ID	Scores
1	30
2	40
3	50
4	30
6	50
7	50
8	40
9	80
10	30
11	50
12	50
13	50
14	50
15	50
16	50
17	50
18	60
19	50
20	60

Note: *From Round 7 the text has changed: 'The total score allocated by you is: … The maximum scores you can allocate is 950. ' The second sentence appeared only where scarcity was introduced.



The score other participants gave you on average: 45.8
Your payoff in this round (in HUF): 1416

Note: *The payoff was reduced/increased only in paid well (PW) treatments.

3. EVALUATING MECHANISMS THAT COULD SUPPORT CREDIBLE REPUTATIONS AND COOPERATION: CROSS-CHECKING AND SOCIAL BONDING[9]

3.1. Introduction

The problem of cooperation has received multidisciplinary attention (see Ostrom et al. 1999, Papadopoulos 2003, Kauser and Shaw 2004, Melis and Semmann 2010 for review) due to its prevalence for a variety of contexts in life. As individual interests work against cooperation, it is a puzzle why cooperation is observed at all, particularly among individuals who are not related to each other and are not engaged in repeated interaction. For such situations, indirect reciprocity has been proposed as a solution (Alexander 1987, Nowak and Sigmund 1998, Milinski, Semmann and Krambeck 2002, Panchanathan and Boyd 2004). It has been suggested that humans have been able to solve the problem of cooperation beyond repeated encounters in small groups because they could rely on informal tools that facilitated the efficiency of downstream indirect reciprocity mechanisms (Nowak and Sigmund 2005, Wu et al. 2020). Gossip is believed to be such an informal tool that enables cooperation as it transmits key information about third parties who are potential interaction partners and hence facilitates the selection of cooperative choice against partners who have good reputation (Nowak 2006, Smith 2010, Milinski 2016, Giardini and Vilone 2016). Gossip may stem from sanctioning motives by which individuals can punish or pose a threat to individuals who were about to exploit cooperation efforts (Gintis 2000, Bowles and Gintis 2004, Fehr and Fischbacher 2003, Molho and Wu 2021). The alleged relationship between gossip and cooperation through the construction of

[9] This chapter was published as: Samu, F. and Takács, K., 2021. Evaluating mechanisms that could support credible reputations and cooperation: cross-checking and social bonding. Philosophical Transactions of the Royal Society B, 376(1838), p.20200302.

reputations has received empirical support in laboratory experiments (Sommerfeld et al. 2007, Sommerfeld et al. 2008, Feinberg et al. 2012, Feinberg et al. 2014 Fonseca and Peters 2018, Samu et al. 2020).

Explanations that link gossip to cooperation are valid only if we assume that gossip contains real information and negative gossip targets those individuals who attempted to exploit cooperation efforts. Gossip, however, is not necessarily honest and credible (Dores Cruz et al. 2021, Fonseca and Peters 2021, Giardini et al. 2021, Hess and Hagen 2021). Distortion might occur from misinterpretation of actions (cf. using first-order social norms, Ohtsuki and Iwasa 2004, Ohtsuki and Iwasa 2006), but it could also be the result of strategic manipulation by the sender (Duffy et al. 2002).

Once gossip is not in line with actions, reputations on which individuals base their decisions become unreliable, so over time, they lose information value. As a consequence, cooperation collapses if it is built up on the shaky ground of miscredited gossip (Roberts 2008, Számadó et al. 2016). Therefore, how gossip could help establish cooperation needs a more thorough investigation. For this purpose, we need to be aware of mechanisms that can maintain the credibility of gossip reputations and we need to know if reliable reputations are sufficient for the maintenance of cooperation. We propose three mechanisms that might be linked to honest gossip, reliable reputations and could undermine or empower cooperation conditional on reputational information.

(a) Competition decreases the reliability of gossip

The transmission of reputational information might not be honest due to the conflict of interest between the sender and the target. Competition for profitable partners (Roberts 1998, Barclay 2004), for social status (Faris and Ennett 2012, Snellman et al. 2019) or for reputation-related benefits (Samu et al. 2020) could all create conflicts of interest. Regardless of the ultimate goal, a good reputation is the target of the competition itself for which both the sender and the target are competing. If reputation is a restricted good, then the conflict of interest might more likely be realized and taken into consideration in communication decisions.

Accordingly, the strive for good reputation drives not only generosity (Barclay and Willer 2007, Piazza and Bering 2008, Barclay 2010, Wu et al. 2016, Bird et al. 2018), but as an alternative tool for individuals to improve their relative rank, also dishonest gossip about rivals (Paine 1967, Barkow 1992, Faris and Felmlee 2011). Unlike random noise (Fehr and Sutter 2019) and exaggeration (Fonseca and Peters 2018), once such strategic misrepresentations are of a realistic possibility, the reliability of social information exchange could be questioned (Mills and Grant 2009) and the alleged link between gossip and cooperation is broken (Peters and Fonseca 2020). In previous experiments, dishonesty was brought about by competition between the sender and receiver of gossip (Peters and Fonseca 2020), but it has not been tested whether people will mislead their audience with dishonest information if they have a conflict of interest only with the target. We investigate how competition for reputational benefit contributes to the greater presence of dishonest gossip signals and indirectly, how this possible strategic misrepresentation affects reputation-based cooperation.

(b) Cross-checking increases the reliability of gossip

Individuals actively seek social information to condition their future actions on a better-informed ground (Swakman et al. 2016). If the same evaluative content is received from multiple sources, then the reliability of gossip increases (Hess and Hagen 2006). As the number of sources increases, dishonesty may be deterred (Giardini and Conte 2012, Boyd and Mathew 2015), since it can be better discovered (Mercier 2012), possibly implying a cost for the sender (De Backer and Gurven 2006). There is no agreement in the literature if multiple sources should be independent in order to channel in information from diverse sources (Harkins and Petty 1987) or they should rather originate from trusted and well-embedded sources from the local network (Burt 2005). It is known, however, that in an unstructured information regime, more gossip better facilitates individual inclinations towards cooperation (Sommerfeld et al. 2008).

Previously, complete information about partners' previous behaviour was condensed in gossip statements and an empirical study on the effect of multiple but uncertain gossip on reputation is still a 'missing piece' (Sommerfeld et al. 2008, p. 2534). In this study, we address this gap by testing the effect of cross-checking by multiple sources on the reliability of gossip.

(c) Social bonding increases the reliability of gossip

Gossip is certainly more than just a form of informal punishment or a deterrence device to avoid free riding. It has been shown that gossip could harmonize the relationship between the sender and the receiver and strengthen their social bonding (Ellwardt et al. 2012). This way, gossip has a similar affiliative impact among humans (Dunbar and Dunbar 1998, Dunbar 2004) as social interactions in other species such as social play (Shimada and Sueur 2018), sensitive touch (Dunbar 2010), food sharing (Wittig et al. 2014), gestural modality (Roberts and Roberts 2017) and grooming (De Waal and Waal 2007, Hemelrijk and Ek 1991, Dunbar 1993), which provide necessary preconditions for cooperation in a situation with conflict, such as mobilization against external or internal threats. More attention to prosocial norms, and mutual expectations about corresponding behaviour, which develop unconsciously as a result of informal communication, can contribute to higher commitment to cooperation (Bicchieri and Lev-On 2007, Torsvik et al. 2011, Przepiorka and Diekmann 2021).

Beyond the role of gossip in unconscious bonding, people can also consciously use gossip to form partnerships (Van de Bunt et al. 2005). We argue that social bonds are created through gossip only if social information is honest, because dishonesty decreases the reputation of the sender (Wilson et al. 2000) and only honest reputational information can lead to a trusted relationship (Bellucci et al. 2019, Bellucci and Park 2020). In this study, we examine the extent to which the two proposed corrective mechanisms (cross-checking and social bonding) can mitigate the potential negative impact of competition.

3.2. Methods

(a) Participants

Two hundred and thirty-four students of the Corvinus University of Budapest participated in a laboratory experiment between January and May 2019. The call was advertised through the university e-mail system and any interested person was able to apply for the experiment through a separate recruitment interface. After arrival to the laboratory, instructions were displayed on participants' screens and were distributed in hard copy as well. Processing of the instructions was tested with questions. Players participated in the experiment anonymously. In order to make participants traceable during the experiment, we identified them with names of planets' moons. All names started with different letters of the alphabet to assist memory capacities. The experiment lasted for an average of 45 min, and it took an average of 10 min to complete the questionnaire following the experiment. The final profit was calculated as the average payoff of six randomly selected rounds. In addition to the final payoff, a show-up fee (HUF 1000) had been paid to the participants. The average payoff was HUF 1807 (approx. 5 EUR). The experiment was programmed with z-Tree (Fischbacher 2007).

(b) Design

We manipulated (1) the level of competition and (2) the presence of mechanisms that can maintain the credibility of gossip (cross-checking, and social bonding) in our experiment between sessions. We introduced competition to increase the likelihood of dishonest gossiping and test whether cross-checking and social bonding mechanisms can eliminate incentivized dishonesty about rivals. Therefore, we interacted manipulation 1 with manipulation 2. With a control condition in which neither cross-checking nor social bonding opportunities were present, we obtained a 2 (competition: high, low) × 3 (mechanism for credible gossip: control, cross-checking, social bonding) factorial design.

Each possible treatment was played in two sessions, so we organized a total of 12 sessions. We had 20 participants in 10 out of the 12 sessions. Eighteen were present in one (low competition—control) and 16 in another session (high competition—cross-checking).

The experiment was divided into two phases. The first phase covered the first five rounds, the second phase lasted from round six until the end of the experiment. Participants did not know when the experiment would end. In the first phase, participants played a Prisoner's

Dilemma game (PD); in the second phase, in addition to the PD, they had the opportunity to gossip and evaluate others (Figure 1).

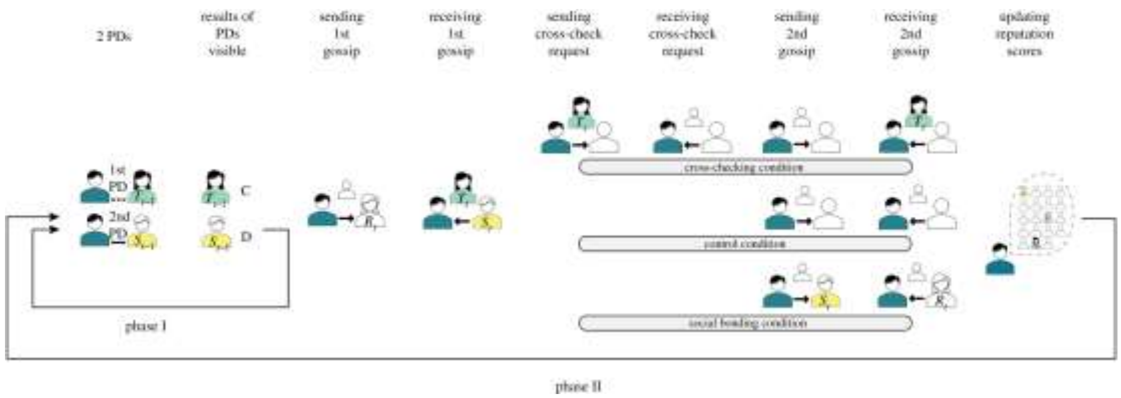


Figure 1. Steps of the experiment within one round. Each round starts with a Prisoner's Dilemma (2 PDs) game, followed by gossip exchange according to treatments and the assessment of other participants' trustworthiness (updating reputation scores).

(c) Procedure

(i) Phase I: basic level of cooperation without communication

At the beginning of each round, individuals were randomly paired with two other players whose fictitious names appeared on the screen and played separate two-person PD games with them (see translated screen 1 in electronic supplementary material, S2). Neutral framing was used in the experiment: options were labelled with letters (*L* and *R*). Outcomes were set as follows. If both players cooperated, they earned HUF 1500 (*R*); in contrast, if both defected, they received HUF 500 (*P*). A person who cooperated while the partner defected was not entitled to payment (*S*). Conversely, the partner's payment was HUF 2000 (*T*). The payoff structure was calibrated such that the index of cooperation (Rapoport A. 1967, Van Lange et al. 2014; $(R - P)/(T - S) = 0.5$) shows a moderate conflict between self- and group interest. Participants had 23 s to decide in the two PD games. If players ran out of time, they got HUF 0, and their PD partner's payoffs depended on their decisions (HUF 0 after cooperation, HUF 500 after defection). In this regard, running out of time was equivalent to defection (cf. Podder et al. 2021), so it could not be used as a costly punishment action. In the first round, 63 players (26.9%) ran out of time, and 44 (18.8%) in the sixth round. Outside these introductory rounds, typically 1–2 people (M = 1.52) did not decide in time. Participants saw the results of their own games on the subsequent screen (see screen 2 in electronic supplementary material, S2).

(ii) Phase II: the reliability of gossip and its effect on cooperation

In the second phase of the experiment, participants played the same PD games as before. In addition, changes were introduced regarding gossip opportunities and reputation building. After the PD game, gossip could be sent to a randomly selected participant (Figure 1 or screen 3 in electronic supplementary material, S2). In each round, participants could send two messages. The fictitious names of gossip targets and receivers were displayed on the screen. Participants could select the valence of gossip from three options indicated by happy, neutral and sad emoticons. We have employed emoticons as they simplify and clarify the content of reputation scores and translate evaluations into positive, neutral or negative judgement. Sending gossip was free and optional and was possible within a limit of 18 s. After the first gossip message, we manipulated how the second message proceeded (Figure 1 and section about manipulation 2).

After the first round in phase II (round 7), players played one of the PDs with their gossip partner from the previous round. The other PD partner was the target of gossip from the previous round. To control who is playing with whom in the next round, the target of the gossip was randomly selected. In one round, only half of the matching resulted in PDs with previous gossip senders and targets. The inverse rule has been applied to the other half of the participants: they played with the receiver of their first gossip and who received a message about them (Figure 2). The computer determined randomly who belongs to which half at the beginning of each round. Players were aware of these matching rules.

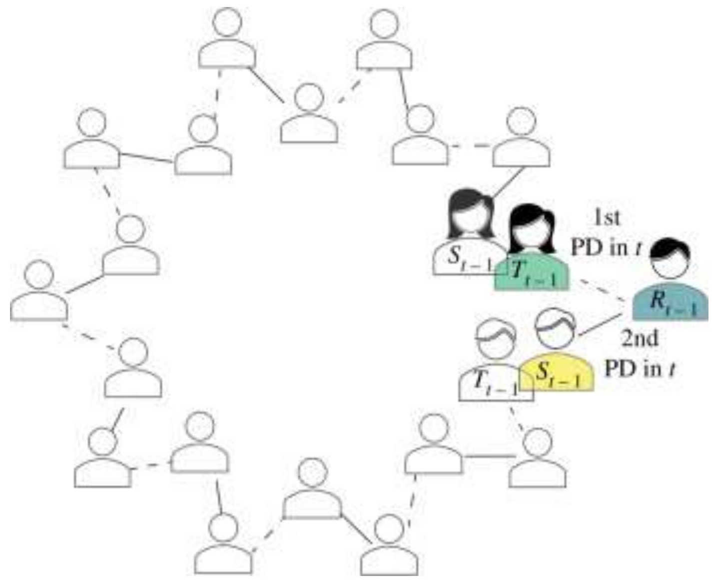


Figure 2. PD partner matching. In the second phase of the experiment, in each round, players were drawn into one of the two roles that determine who they play the PD game in that round with: (i) half of the participants (outer circle, R) played one PD with their first gossip source in the previous round (St−1) and one PD with a target of the gossip from this source (Tt−1); and (ii) the other half of the participants (inner circle, S and T), accordingly, played one PD (solid line) with a receiver of the gossip sent by them (Rt−1, for St−1, not tagged for Tt−1) and one PD (dashed line) with a participant who received gossip about them (Rt−1, for Tt−1, not tagged for St−1).

From round 6, besides gossip, players could assign reputation scores to other participants. They were asked to evaluate on a scale of 0–100 according to how much they 'trust other participants'. These individually assigned private reputation scores are hence not consensual. In round 6, everyone's score was set to a starting value of 50, but changes were saved to subsequent rounds, thus, players were able to use the saved reputation scores they assigned. Fifty seconds were available for the assignment of reputation scores. Each round ended with a summary where players learned their own average reputation scores and those of their rivals, as well as their adjusted payoffs in the given round.

(iii) Manipulation 1: competition for reputation

Above the PDs, reputation scores played a role for the payoffs in phase II. Payoffs were adjusted according to the reputation score players received on average relative to a reference group of five participants (rivals). By introducing small rival groups, we tested whether players try to wreck rivals' reputation by dishonest negative gossip. Rivals were selected randomly at the very beginning of phase II, and they remained the same until the end of the experiment.

A deviance of the participant's mean reputation score relative to the rivals' decreased/increased the participant's payoff. Payoffs from the PDs have not been altered for those who received the same score on average as their rivals. The magnitude of the alteration was determined by the strength of competition (high versus low). One-unit deviance reduced/increased payoffs by HUF 20 (approx. 5.5 euro cents) in high competition and by HUF 2 (approx. 0.55 euro cents) in low competition. Thus, manipulation 1 modified the strength of the competition for reputation scores.

(iv) Manipulation 2: mechanisms that can maintain the credibility of gossip

Cross-checking. In the cross-checking condition, we allowed players to ask for a second gossip about the same target (see the top row in Figure 1). Cross-checking gossip about potential partners could lead to a more reliable assessment of others' willingness to cooperate. In the control condition, the second gossip could be applied to a new target.

Social bonding. In the social bonding condition, we manipulated whether gossip could be reciprocated. We analysed the effect of this affiliative action on the reliability of gossip, reputations and cooperation. In each round, players could send two messages in a row to a pre-designated receiver. In the social bonding manipulation, the second gossip could be reciprocated to the source of the first gossip (see the bottom row in Figure 1). In the control treatment, the receiver of the second message was a new subject (see the middle row in figure 1). We consider this reciprocated action as a less costly opportunity for bonding before participants face a more conflicted situation in the next round's PD game (see matching of next PD partners in Figure 2).

3.3. Results

(a) Descriptive statistics

(i) Cooperation

Baseline cooperation without communication in the first five rounds (38.7%) has increased in round 6, after the introduction of gossip and the opportunity for reputation building (52.1%). Afterwards, cooperation eroded gradually till the last round of the experiment (29.6%). High competition induced an average level of 43.7% cooperation, while the cooperation rate in the low competition was 31.7%. Cross-checking generated an average cooperation rate of 30.9%, while social bonding produced an average cooperation rate of 40.3% similar to the control condition (41.8%), in which neither social bonding nor cross-checking opportunities were present (Figure 3).

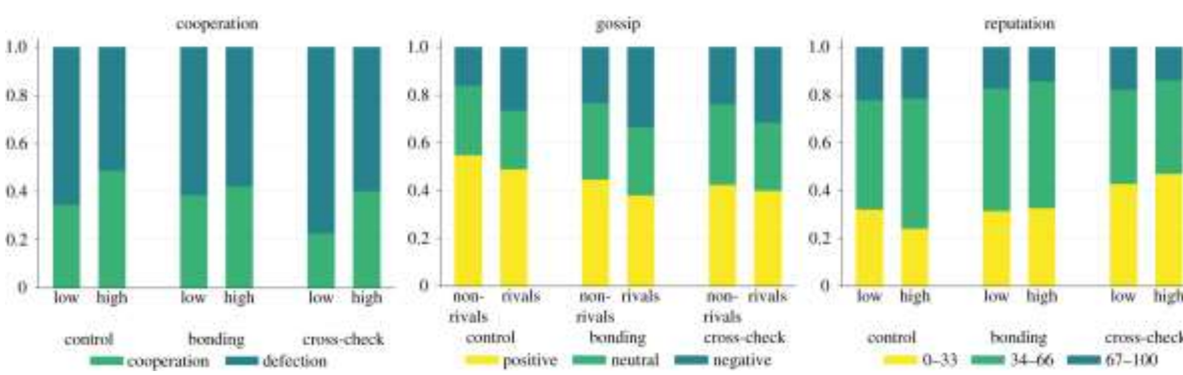


Figure 3 Cooperation, the valence of the gossip and trust by manipulations. Cooperation is higher in high-competition treatments. Negative gossip is more prevalent when rivals are the targets of gossip. Reputation scores are lower in cross-checking treatments.

(ii) Gossip

Participants used gossip in 86.7% of their opportunities, both under high (87.0%) and low competition (86.5%), but the exploitation of gossip opportunity varied by treatment conditions (control: 95.4%; social bonding: 90.3%; cross-checking: 74.1%). In both the social bonding and cross-checking conditions, opportunities to send gossip were limited because they depended on the initiation of the gossip partner. In the social bonding condition, a second gossip could only be sent in response to the first gossip if it had been sent (in 91.1% of cases, participants used the first gossip opportunity). In the cross-checking treatment, participants could only send the second gossip if they received a request for cross-checking (68.4% of possible requests were sent) given that there was any first gossip to verify (the first gossip was sent in 91.8% of cases). On average, gossip was mainly positive (46.3%), less often neutral (30.2%) or negative (23.4%).

(iii) Reputation

Reputation scores (measured on a 100-point scale) did not differ considerably between treatments with low (42.2) and high competition (42.6). Reputation scores reached their lowest average value in the cross-checking (37.5) and in the social bonding condition (41.8), while the average value in the control condition was slightly lower than the initial score of 50 (47.9). In the following sections, we provide insights into the micro-level mechanisms that are responsible for these patterns at the macro level.

(b) Multilevel mixed-effects models

For the establishment of reputation-based cooperation mediated by second-hand information, such as gossip, three associations are quintessential. First, gossip should be honest, such that it reflects past behaviour. Second, gossip should be believed by the receiver and incorporated into the receivers' perception of the target. Third, receivers have to make decisions according to this cognitive image when they decide about cooperation or defection against the target.

Competition can induce distortion in the first step by encouraging dishonest gossip about rivals. This can make the entire reputation system unreliable because the distortion impedes subsequent associations. If second-hand information or bonding considerations between the sender and the receiver do not provide guidance to make appropriate decisions, individuals will rather choose defection as a secure strategy that may result in the collapse of cooperation. In the following, we examine the presence of dishonesty, its potential escalation by competition, and whether social bonding and cross-checking can correct this distortion. Applying mixed effect multilevel models, we adjust our analysis to individual's repeatedly observed decisions.

(i) Reliability of gossip

Using multilevel ordered logistic models, we found that, regardless of all other factors, gossip about rivals was more negative ($\beta = -0.29$, $p < 0.001$, model 1; electronic supplementary material, table S1) suggesting that players tried to improve their own position to the detriment of rivals. When the competition was low, in the cross-checking condition, we did not detect any dishonesty about rivals ($\beta = 0.42$, $p < 0.05$, model 2; electronic supplementary material, table S1), which means that the opportunity for cross-checking significantly holds back negative gossip about rivals compared to the control condition. This apparent counterforce disappeared in high competition ($\beta = -0.86$, $p < 0.01$, model 2; electronic supplementary material, table S1), despite the fact that dishonesty has not been intensified by competition ($\beta = 0.27$, $p = 0.18$, model 2; electronic supplementary material, Table S1).

Apart from the distortion created by rivalry, gossip was sent in an honest way in the sense that it was aligned with targets' PD decisions (if sender was an involved PD partner in a given round): if the target defected, then gossip was less positive ($\beta = -1.36$, $p < 0.001$, model 2; electronic supplementary material, table S1); while if the target of gossip cooperated with the sender, then gossip was more positive ($\beta = 1.35$, $p < 0.001$, model 2; electronic supplementary material, Table S1).

Since senders were not always in direct encounters with gossip targets, gossip could rest on players’ private reputation assessment as well. The higher the target's reputation was, the more likely a positive message was sent about that person ($\beta = 0.03$, $p < 0.001$, model 2; electronic supplementary material, table S1). Compared to the control group, in the cross-checking treatment, the likelihood of sending positive gossip increased less steeply as the reputation score increased ($\beta = -0.01$, $p < 0.001$, model 2; electronic supplementary material, Table S2). In other words, gossip about players with good reputations was less positive in this treatment.

(ii) Building a reputation system on believed information

Being aware of the presence of dishonesty, we examine whether gossip was believed and was incorporated into private reputation assessments. Participants privately assigned reputation scores to others, to preserve and be able to recall their previous behaviours. When doing so, they potentially integrated evaluations received from others into their scores. Participants modified their evaluations in line with the gossip they received. Positive messages increased ($\beta = 7.42$, $p < 0.001$, model 1; electronic supplementary material, table S3), negative messages decreased ($\beta = -5.14$, $p < 0.001$, model 1; electronic supplementary material, table S3) the allocated reputation scores to the target compared to those about whom neutral gossip have been heard. There were differences between the manipulations with regard to how messages had been incorporated into reputation ratings. In high competition, negative messages decreased reputations with a larger magnitude ($\beta = -2.53$, $p < 0.05$, model 3; electronic supplementary material, table S3) and positive messages were less rewarding ($\beta = -2.49$, $p < 0.01$, model 3; electronic supplementary material, table S3). In the social bonding condition, positive messages increased targets' reputation scores more than in the control condition ($\beta = 4.67$, $p < 0.001$, model 2; electronic supplementary material, table S3).

We note that the trustworthiness of the gossip source played a role in accepting gossip as true. No credit was given to negative messages when the source of gossip had a bad reputation ($\beta = -2.22$, $p = 0.06$, model 2; electronic supplementary material, table S4). Moreover, the penalty for negative gossip increased as the reputation of the sender improved ($\beta = -0.05$, $p < 0.05$, model 2; electronic supplementary material, table S4). Besides gossip, as expected, reputations were formed by participants' direct experience as an involved party in the PD: assigned reputation scores were adjusted in the positive direction after cooperation ($\beta = 8.63$, $p < 0.001$, model 1; electronic supplementary material, table S3), and in the negative direction after defection by the interaction partner (β

= −9.06, $p < 0.001$, model 1; electronic supplementary material, table S3).

Apart from first- and second-hand information, two other factors affected participants' assessments. Participants appreciated the gossip they received: gossip senders received slightly better reputation scores ($\beta = 0.59$, $p < 0.05$, model 1; electronic supplementary material, table S3), and those who could gossip but did not send any messages received lower ratings ($\beta = -3.86$, $p < 0.001$, model 1; electronic supplementary material, table S3). Also, reputation scores assigned to rivals were significantly lower ($\beta = -2.05$, $p < 0.001$, model 1; electronic supplementary material, table S3), even if scores from rivals did not affect individuals' payoff.

(iii) Reputation-based cooperation

Regarding the third link of the main narrative, we found evidence that cooperation was conditional on the reputation scores of PD partners ($\beta = 0.02$, $p < 0.001$, model 1; electronic supplementary material, table S5). From the manipulations, only high competition led to a higher level of cooperation regardless of the partner's reputation (participants cooperated more even if their partners had a bad reputation; $\beta = 1.42$, $p < 0.001$, model 2; electronic supplementary material, table S5), but the positive impact of reputation scores was weaker in this treatment ($\beta = -0.01$, $p < 0.001$, model 2; electronic supplementary material, table S5) and the likelihood of cooperation with trustworthy individuals returned to the level of treatments with low competition. The positive effect of strong competition kept the otherwise declining cooperation ($\beta = -0.11$, $p < 0.001$, model 2; electronic supplementary material, table S6) at a higher level over time ($\beta = 0.06$, $p < 0.001$, model 2; electronic supplementary material, table S6).

(iv) Overall reflectivity

Finally, we provide an overview of whether a reliable reputation system has been established by honest gossip, gossip-based trust formation and reputation-based cooperation. As a result of these links, a reliable reputation system can develop that reflects past actions well; thus it provides a good guide for individuals to conditionally cooperate.

Surprisingly, despite dishonest gossip about rivals, the reputation system helped subjects to make good decisions in each condition: the more someone cooperated in previous rounds, the more likely others cooperated with that person ($\beta = 1.25$, $p < 0.001$, model 1; electronic supplementary material, table S7). The overall association did not differ between conditions (see models 2, 3, 4; electronic supplementary material, table S7). Even if we see differences in the strengths of the operating mechanisms between conditions, we observed a good overall efficiency of the reputation system in our experiment.

3.4. Conclusion

A reputation system can effectively maintain cooperation only if it is based on reliable information spreading. Gossip—an evaluative communication about third parties—could be the channel of reliable information transmission and hence could contribute to the maintenance of cooperation (Wu et al. 2016 for review). There is significant doubt, however, about why gossip should be honest and reliable at all (Smith 2014). In this study, we investigated mechanisms that could alter whether gossip could be a successful informal mechanism that establishes cooperation through the construction of reliable reputations.

First, we argued that strong direct rivalry for reputations could increase opportunistic use of gossip and hence decrease the reliability of the information received. We have designed the high-competition condition in our experiment in a way that direct rivalry with a set of other participants meant a distribution of monetary payoffs depending on *relative* reputations. Second, we argued that once the opportunity is given, individuals actively seek and cross-check social information to condition their future actions on a better-informed ground, which improves the reliability of reputations they assign to others. While not just sending, but also seeking gossip possibly takes place in complex ways in human interactions, we implemented cross-checking as a single opportunity to ask a second opinion about the same target. Third, we argued that social bonding motives could increase the credibility of social information exchange and hence make reputations reliable. Although it was not possible to create real social bonds between participants in the experiment, we selected a single characteristic that is typical of social bonding and friendship formation and could also be introduced in an abstract experimental setting: *reciprocity in communication*. Note that reciprocity in communication did not mean reciprocity in interactions as participants played PD games against different partners to follow the settings described in models of indirect reciprocity (Alexander 1987, Nowak and Sigmund 1998, Milinski et al. 2002, Panchanathan and Boyd 2004). We expected that both cross-checking and social bonding operationalized as reciprocity in communication between the sender and the receiver could be efficient mechanisms ensuring honesty of gossip in conditions of intense competition for reputations.

Even if gossip and reputation scores were mutually aligned with each other and with the PD decisions, cooperation did not emerge to a very high rate in any of the conditions. Competition for reputations had divergent effects in our experiment. On the one hand, messages about rivals were more negative, which diminished the reliability of assigned reputations. On the other hand, cooperation was affected positively by the strength of competition. In line with competitive helping theory, rivalry increased cooperation regardless of the reputation of partners (Roberts et al. 2021). Still, no escalation of cooperation was observed; only the decline of cooperation slowed down (cf. Fischbacher et al. 2001).

Though reputation scores grew more as a result of positive messages received, the possibility of social bonding did not cause significant improvement for cooperation. Our results are consistent with the fact that people place more weight on positive information if it comes from a stronger social bond (Bozoyan and Vogt 2016). The integration of received information from trusted sources is important for a well-functioning reputation system, but as social bonding did not improve significantly how reputations are used to condition behaviour, this treatment did not substantially improve cooperation overall.

In the cross-checking condition, we observed a greater cautiousness of participants. Participants were less courageous in sending positive gossip about trustworthy partners. Besides greater cautiousness, participants often received conflicting information about the same target (see electronic supplementary material, table S8), which may lower the reliability of communication even compared to no information (Kuttler et al. 2002). Mixed gossip could have an averaging (Sommerfeld et al. 2008) and a majority effect (Laidre et al.

2013) on reputations. Surprisingly, people inclined to doubt multiple negative opinions as well (Sommerfeld et al. 2008, Hess and Hagen 2006) (see β = 0.96, p = 0.62, model 1; electronic supplementary material, table S9).

Participants in the cross-checking and social bonding conditions were assigned lower reputation scores in general. Lower reputation scores in these conditions—measured as trustworthiness—may have been caused by a general lack of trust caused by the inefficient (Zand 1972) and sometimes contradicting information participants received. Social information needs to be available in large amounts to assist cooperation (Giardini and Vilone 2016, Romano et al. 2021). Correspondingly, the reputation of gossip sources was eroded if they failed to provide information.

Confidence in gossip from trustworthy sources was higher (Kuttler et al. 2002, Pasquini et al. 2007). People seek information from sources considered trusted (Van de Bunt et al. 2005), probably because of their (perceived) good access to information. Therefore, gossip and the dynamics of reputation and cooperation should be considered from the perspective of the social network structure and the position of relevant individuals within (see Takács et al. 2021 for review, Dumas et al. 2021).

Our results suggest that a reliable reputation system is not a sufficient condition for cooperation in situations of moderate conflict of interest. At the same time, we found that relative competition seems to play an important role for cooperation, which could be linked with keeping up with others (loss avoidance) or achieving reputational benefits (status maximization) for the development of widespread human cooperation (Roberts 1998, Barclay 2004, Barclay and Willer 2007, Wu et al. 2016, Herrmann et al. 2021, Raihani and Smith 2015, Raihani and Bshary 2015).

Overall, while we found effects of intensified competition, cross-checking and social bonding for the reliability of gossip, building up of reputations, and partly on conditional behaviour, none of these mechanisms in their abstract form and out of social context were able to sustain a high level of cooperation in the laboratory. Note that gossip was implemented in a very simplified form, as transmission of evaluative social information (sending an emoticon) about the target. This certainly limits the generalizability of our results to empirical situations in which the power of gossip is enhanced in extensive communication.

3.5. Supplementary

S1 Supplementary tables and figures

Table S1. Multilevel Mixed-Effects Ordered Logistic Regression – Conditional Effects of Rivalry on Gossip Choices; Simple Effects

Dependent variable: gossip choice (1: ☹, 2: 😐, 3: 😉)	Model 1	Model 2
Rival (0: target is not sender's rival, 1: target is sender's rival)	-0.29***	-0.40**
	[0.06]	[0.14]
Manipulation 1 (baseline: low competition)		
high competition	0.06	0.22
	[0.20]	[0.35]
Manipulation 2 (baseline: control)		
social bonding	-0.39	-0.19
	[0.25]	[0.35]
cross-checking	-0.32	-0.3
	[0.25]	[0.35]
Manipulation 1 – Rival interaction		
high competition # rival		0.27
		[0.20]
Manipulation 2 – Rival interaction		
social bonding # rival		0.13
		[0.19]
cross-checking # rival		0.42*
		[0.21]
Manipulation 1 – Manipulation 2 interaction		
high competition # social bonding		-0.35
		[0.50]
high competition # cross-checking		-0.02
		[0.51]
Manipulation 1 – Manipulation 2 – Rival interaction		
high competition # social bonding # rival		-0.4
		[0.27]
high competition # cross-checking # rival		-0.86**
		[0.29]
Target's reputation score before update (0-100)	0.03***	0.03***
	[0.00]	[0.00]
Target's PD decision in the same round (baseline: not a PD partner)		
cooperate	1.34***	1.35***
	[0.15]	[0.15]
defect	-1.36***	-1.36***

	[0.11]	[0.11]
Round	-0.02***	-0.02***
	[0.00]	[0.00]
cutoff1	-1.09***	-1.00***
	[0.22]	[0.27]
cutoff2	0.91***	1.00***
	[0.22]	[0.27]
random intercept variance between subjects	2.20***	2.19***
	[0.25]	[0.25]
N of decisions	7307	7307
N of participants	234	234

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; standard errors in brackets; ††random intercept variance within subjects has been fixed to $\pi^2/3=3.29$.

Table S2. Multilevel Mixed-Effects Ordered Logistic Regression – Conditional Effects of Reputation Scores on Gossip Choices; Simple Effects

Dependent variable: gossip choice (1: 😟, 2: 😐, 3: 😊)	Model 1	Model 2	Model 3 – Low competition	Model 4 – High competition
Target’s reputation score before update (0-100)	0.03***	0.03***	0.04***	0.03***
	[0.00]	[0.00]	[0.00]	[0.00]
Manipulation 1 (baseline: low competition)				
high competition	0.49	0.05		
	[0.38]	[0.20]		
Manipulation 2 (baseline: control)				
social bonding	-0.37	-0.52	-0.4	-0.65
	[0.38]	[0.27]	[0.40]	[0.36]
cross-checking	0.4	0.11	0.39	-0.18
	[0.38]	[0.27]	[0.40]	[0.37]
Manipulation 1 – Manipulation 2 interaction				
high competition # social bonding	-0.3			
	[0.53]			
high competition # cross-checking	-0.6			
	[0.54]			
Manipulation 1 – Target’s reputation score before update interaction				
high competition # reputation	0			
	[0.00]			
Manipulation 2 – Target’s reputation score before update interaction				
social bonding # reputation	0.01	0	0.01	0
	[0.00]	[0.00]	[0.00]	[0.00]
cross-checking # reputation	-0.02***	-0.01***	-0.02***	-0.01
	[0.00]	[0.00]	[0.00]	[0.00]
Manipulation 1 – Manipulation 2 – Target’s reputation score before update interaction				
high competition # social bonding # reputation	0			
	[0.00]			
high competition # cross-checking # reputation	0.01			
	[0.01]			
Rival (baseline: target is not sender’s rival)				
target is sender’s rival	-0.29***	-0.29***	-0.22**	-0.36***
	[0.06]	[0.06]	[0.08]	[0.08]
Target’s PD decision in the same round (baseline: not a PD partner)				
cooperate	1.36***	1.36***	1.11***	1.57***
	[0.15]	[0.15]	[0.22]	[0.21]
defect	-1.36***	-1.36***	-1.41***	-1.34***
	[0.11]	[0.11]	[0.15]	[0.16]
Round	-0.02***	-0.02***	-0.01	-0.03***
	[0.01]	[0.01]	[0.01]	[0.01]

cutoff1	-0.83**	-1.03***	-0.75*	-1.40***
	[0.28]	[0.23]	[0.31]	[0.29]
cutoff2	1.19***	0.98***	1.45***	0.45
	[0.28]	[0.23]	[0.31]	[0.28]
random intercept variance between subjects	2.22***	2.22***	2.48***	2.00***
	[0.25]	[0.25]	[0.40]	[0.33]
N of decisions	7307	7307	3674	3633
N of participants	234	234	118	116

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; standard errors in brackets; ††random intercept variance within subjects has been fixed to $\pi^2/3=3.29$.

Table S3. Multilevel Mixed-Effects Linear Regression – Conditional Effect of Gossip on Reputation Assignment; Simple Effects

Dependent variable: How much subject trust others (0-100)	Model 1	Model 2	Model 3	Model 4
Received Gossip (baseline: 😐)†				
no gossip about others	2.23***	2.21**	2.50***	1.91
	[0.41]	[0.71]	[0.57]	[1.01]
😊	7.42***	6.35***	8.66***	6.19***
	[0.52]	[0.87]	[0.73]	[1.25]
☹	-5.14***	-6.40***	-3.88***	-5.25***
	[0.62]	[1.10]	[0.88]	[1.55]
Manipulation 1 (baseline: low competition)				
high competition	-0.25	-0.25	0.43	-0.43
	[0.82]	[0.82]	[1.15]	[1.99]
Manipulation 2 (baseline: control)				
social bonding	-2.04*	-3.05*	-2.04*	-3.15
	[1.00]	[1.39]	[1.00]	[1.96]
cross-checking	-3.16**	-2.3	-3.16**	-3.51
	[1.02]	[1.42]	[1.02]	[2.00]
Manipulation 1 – Manipulation 2 interaction				
high competition # social bonding				0.21
				[2.78]
high competition # cross-checking				2.47
				[2.85]
Manipulation 1 – Received Gossip interaction				
high competition # no gossip about others			-0.58	0.6
			[0.82]	[1.43]
high competition # 😊			-2.49*	0.34
			[1.04]	[1.74]
high competition # ☹			-2.53*	-2.34
			[1.24]	[2.20]
Manipulation 2 – Received Gossip interaction				
social bonding # no gossip about others		0.88		1.1
		[0.99]		[1.38]
social bonding # 😊		4.67***		6.81***
		[1.24]		[1.74]
social bonding # ☹		0.78		1.26
		[1.49]		[2.10]
cross-checking # no gossip about others		-0.97		0.6
		[1.03]		[1.44]
cross-checking # 😊		-1.71		0.25
		[1.31]		[1.84]
cross-checking # ☹		3.58*		3.09
		[1.59]		[2.24]

Manipulation 1 – Manipulation 2 – Received Gossip interaction				
high competition # social bonding # no gossip about others				-0.45
				[1.98]
high competition # social bonding # 😊				-4.52
				[2.49]
high competition # social bonding # ☹				-0.83
				[2.98]
high competition # cross-checking # no gossip about others				-3.19
				[2.05]
high competition # cross-checking # 😊				-4.06
				[2.62]
high competition # cross-checking # ☹				0.98
				[3.18]
Round	-0.14***	-0.14***	-0.14***	-0.14***
	[0.01]	[0.01]	[0.01]	[0.01]
Other Players' Reputation Score in the Previous Round	0.69***	0.69***	0.69***	0.69***
	[0.00]	[0.00]	[0.00]	[0.00]
PD Decision (baseline: not a PD partner)				
cooperate	8.63***	8.64***	8.63***	8.63***
	[0.35]	[0.35]	[0.35]	[0.35]
defect	-9.06***	-9.06***	-9.06***	-9.06***
	[0.27]	[0.27]	[0.27]	[0.27]
Rival (baseline: not rival)				
rival	-2.05***	-2.05***	-2.05***	-2.05***
	[0.15]	[0.15]	[0.15]	[0.15]
Gossip Sender				
sent messages	0.59*	0.58*	0.59*	0.59*
	[0.24]	[0.24]	[0.24]	[0.24]
did not sent any messages	-3.86***	-3.84***	-3.86***	-3.85***
	[0.73]	[0.73]	[0.73]	[0.73]
constant	15.20***	15.29***	14.87***	15.38***
	[0.95]	[1.11]	[1.03]	[1.44]
random intercept variance between subjects	1.83***	1.83***	1.83***	1.83***
	[0.05]	[0.05]	[0.05]	[0.05]
random intercept variance within subjects	2.95***	2.95***	2.95***	2.95***
	[0.00]	[0.00]	[0.00]	[0.00]
N of decisions	79511	79511	79511	79511
N of participants	234	234	234	234

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, standard errors in brackets; †Mixed gossip about the same target was coded as follows: 😊😐→😊, ☹😐→☹, 😊☹→😐

Table S4. Multilevel Mixed-Effects Linear Regression – Conditional Effect of Gossip Sender's reputation on Reputation Assignment to the Target of Gossip; Simple Effects

Dependent variable: How much subject trust others (0-100)		Model 1	Model 2
Target's Reputation Score in the Previous Round		0.61***	0.61***
		[0.01]	[0.01]
PD Decision (baseline: not a PD partner)			
	cooperate	9.21***	9.15***
		[1.33]	[1.33]
	defect	-8.77***	-8.75***
		[1.04]	[1.04]
Rival (0: no, 1: yes)		-2.56***	-2.53***
		[0.59]	[0.59]
Received Gossip (baseline: 😐) †	😊	7.93***	7.60***
		[0.62]	[1.07]
	☹	-4.26***	-2.22
		[0.73]	[1.20]
Manipulation 1 (0: low competition, 1: high competition)		-1.21	-1.21
		[1.12]	[1.12]
Manipulation 2 (baseline: control)			
	social bonding	-1.38	-1.36
		[1.35]	[1.35]
	cross-checking	-2.75*	-2.74*
		[1.39]	[1.39]
Round		-0.18***	-0.18***
		[0.05]	[0.05]
Sender's Reputation Score			0.01
			[0.02]
Sender's Reputation Score – Received Gossip interaction	😊		0.01
			[0.02]
	☹		-0.05*
			[0.02]
	constant	18.85***	18.37***
		[1.52]	[1.67]
	random intercept variance between subjects	2.02***	2.02***
		[0.06]	[0.06]
	random intercept variance within subjects	3.10***	3.10***
		[0.01]	[0.01]
N of decisions		7304	7304
N of participants		234	234

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, †Mixed gossip about the same target was coded as follows: 😊😐→😊, ☹😐→☹, 😊😊→😐

Table S5. Multilevel mixed-effects logistic regression – Conditional Effect of PD Partner's Reputation on PD decisions; Simple Effects

Dependent variable: PD choice (0: defect, 1: cooperate)		Model 1	Model2
Partners Reputation Score		0.02***	0.03***
		[0.00]	[0.00]
Manipulation 1 (baseline: low competition)			
	high competition	0.94**	1.42***
		[0.30]	[0.32]
Manipulation 2 (baseline: control)			
	social bonding	-0.05	-0.05
		[0.36]	[0.36]
	cross-checking	-0.64	-0.65
		[0.37]	[0.37]
Manipulation 1 – Partners Reputation Score interaction			
	high competition # reputation		-0.01***
			[0.00]
Relative Reputation Score†		-0.04	-0.04
		[0.02]	[0.02]
Relative Reputation Score squared		0	0
		[0.00]	[0.00]
Player Played PD as a Receiver		0.08	0.08
		[0.06]	[0.06]
Round		-0.06***	-0.07***
		[0.01]	[0.01]
	constant	-1.06*	-1.28**
		[0.43]	[0.44]
	random intercept variance between subjects ††	5.17***	5.15***
		[0.65]	[0.64]
N of decisions		8492	8492
N of participants		234	234

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, standard errors in brackets; †Original scale contains integers (-24 to +24) and was converted to natural numbers (1 to 48); ††random intercept variance within subjects has been fixed to $\pi^2/3=3.29$.

Table S6. Multilevel mixed-effects logistic regression – Conditional Effect of Round on PD decisions; Simple Effects

Dependent variable: PD choice (0: defect, 1: cooperate)	Model 1	Model 2	Model 3
Round	-0.04***	-0.11***	-0.07***
	[0.01]	[0.01]	[0.02]
Manipulation 1 (baseline: low competition)			
high competition	0.90**	0.06	0.46
	[0.29]	[0.33]	[0.58]
Manipulation 2 (baseline: control)			
social bonding	0.58	-0.2	0.79
	[0.41]	[0.36]	[0.58]
cross-checking	-0.38	-0.86*	0.06
	[0.41]	[0.36]	[0.58]
Manipulation 1 – Manipulation 2 interaction			
social bonding # high competition			-0.43
			[0.82]
cross-checking # high competition			-0.71
			[0.82]
Manipulation 1 – Round interaction			
high competition # round		0.06***	0.05*
		[0.01]	[0.02]
Manipulation 2 – Round interaction			
social bonding # round	-0.06***		-0.03
	[0.01]		[0.02]
cross-checking # round	-0.03*		-0.08***
	[0.01]		[0.02]
Manipulation 1 – Manipulation 2 * Round interaction			
high competition # social bonding # round			-0.04
			[0.03]
high competition # cross-checking # round			0.08**
			[0.03]
Player Played PD as a Receiver	0.1	0.1	0.1
	[0.06]	[0.06]	[0.06]
constant	-0.58	0.28	-0.36
	[0.33]	[0.32]	[0.42]
random intercept variance between subjects	4.99***	4.99***	4.96***
	[0.62]	[0.63]	[0.62]
N of decisions	8404	8404	8404
N of participants	234	234	234

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, standard errors in brackets; † random intercept variance within subjects has been fixed to $\pi^2/3=3.29$.

Table S7. Multilevel mixed-effects logistic regression – Conditional Effect of PD Partner's Previous Cooperativeness on PD decisions; Simple Effects

Dependent variable: PD choice (0: defect, 1: cooperate)	Model 1	Model 2	Model 3	Model 4

PD Partner's Previous Cooperation Rate	1.25***	1.22***	0.73*	0.29
	[0.17]	[0.24]	[0.32]	[0.47]
Manipulation 1 (baseline: low competition)				
high competition	0.79**	0.76*	0.80**	0.7
	[0.28]	[0.32]	[0.28]	[0.57]
Manipulation 2 (baseline: control)				
social bonding	-0.24	-0.24	-0.65	-0.34
	[0.34]	[0.34]	[0.40]	[0.56]
cross-checking	-0.71*	-0.71*	-0.98*	-1.30*
	[0.35]	[0.35]	[0.39]	[0.54]
Manipulation 1 – Manipulation 2 interaction				
high competition # social bonding				-0.6
				[0.80]
high competition # cross-checking				0.69
				[0.79]
Manipulation 1 – PD Partner's Previous Cooperation Rate interaction				
		0.06		0.8
		[0.34]		[0.64]
Manipulation 2 – PD Partner's Previous Cooperation Rate interaction				
social bonding # past behaviour			0.85*†	1.31*†
			[0.43]	[0.62]
cross-checking # past behaviour			0.59	1.15
			[0.43]	[0.62]
Manipulation 1 * Manipulation 2 * PD Partner's Previous Cooperation Rate interaction				
high competition # social bonding # past behaviour				-0.81
				[0.87]
high competition # cross-checking # past behaviour				-1.06
				[0.86]
Player Played PD as a Receiver	0.1	0.1	0.1	0.1
	[0.06]	[0.06]	[0.06]	[0.06]
constant	-1.71***	-1.69***	-1.48***	-1.43***
	[0.30]	[0.30]	[0.32]	[0.40]
random intercept variance between subjects	4.58***	4.57***	4.61***	4.54***
	[0.58]	[0.58]	[0.58]	[0.57]
N of decisions	8404	8404	8404	8404
N of participants	234	234	234	234

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, standard errors in brackets; † it is only significant in one experimental session, †† random intercept variance within subjects has been fixed to $\pi^2/3=3.29$.

Table S8. Distribution of Mixed Gossip about the Same Target in the Cross-Checking Treatment

First Gossip	Second Gossip	N	%
😊	😊	132	17
😊	😐	200	26

😊	😟	180	23
😐	😐	82	11
😐	😟	123	16
😟	😟	54	7

Table S9. Multilevel Mixed-Effects Linear Regression – Reputation Assignment in the Cross-Checking Treatment

Dependent variable: How much subject trust others (0-100)	Model1	Model 2 Low competition	Model 3 High competition
Targets' Reputation Score in the Previous Round	0.56***	0.61***	0.49***
	[0.02]	[0.03]	[0.03]
Received Gossip (baseline: 😐😐)			
😊 😊	8.08***	8.88***	7.62**
	[1.59]	[2.11]	[2.38]
😊 😐	2.4	2.5	2.95
	[1.65]	[2.15]	[2.54]
😊 🙁	1.5	-1.27	4.48
	[1.70]	[2.32]	[2.52]
🙁 😐	-5.19**	-6.43**	-3.14
	[1.87]	[2.46]	[2.85]
🙁 🙁	0.96	0.51	2.01
	[1.91]	[2.63]	[2.79]
Manipulation 1 (0: low intensity, 1: high intensity)	0.12		
	[2.60]		
Round	-0.35***	-0.12	-0.55***
	[0.10]	[0.14]	[0.14]
Rival (0: no, 1: yes)	-2.91**	-0.55	-5.11***
	[1.09]	[1.51]	[1.55]
PD Decision (baseline: not a PD partner)			
cooperate	9.58***	18.91***	3.04
	[2.73]	[4.06]	[3.67]
defect	-5.88***	-4.58*	-8.02**
	[1.76]	[2.27]	[2.79]
constant	20.00***	14.19***	25.67***
	[2.68]	[3.24]	[3.55]
random intercept variance between subjects	2.35***	2.33***	2.37***
	[0.10]	[0.15]	[0.15]
	3.06***	3.04***	3.07***
random intercept variance within subjects	[0.02]	[0.02]	[0.02]
N of decisions	2028	1053	975
N of participants	76	40	36

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, standard errors in brackets; † random intercept variance within subjects has been fixed to $\pi^2/3=3.29$.

Table S10. Multilevel Mixed-Effects Ordered Logistic Regression – Gossip Choices, Conditional Effects of Rivalry; Interaction Effects

Dependent variable: gossip choice (1: 😟, 2: 😐, 3: 😊)		Model 2
Rival (0: target is not sender's rival, 1: target is sender's rival)		-0.63**
		[0.22]
Manipulation 1 (baseline: low competition)		
	high competition	0.12
		[0.54]
Manipulation 2 (baseline: control)		
	social bonding	-0.38
		[0.28]
	cross-checking	-0.3
		[0.35]
Manipulation 1 * Manipulation 2		-0.01
		[0.25]
Manipulation 1 * Rival		0.71*
		[0.31]
Manipulation 2 * Rival		0.21*
		[0.10]
Manipulation 1 * Manipulation 2 * Rival		-0.43**
		[0.15]
Target's reputation score before update (0-100)		0.03***
		[0.00]
Target's PD decision in the same round (baseline: not a PD partner)		
	cooperate	1.35***
		[0.15]
	defect	-1.36***
		[0.11]
Round		-0.02***
		[0.00]
	cutoff1	-1.07***
		[0.26]
	cutoff2	0.94***
		[0.26]
	random intercept variance between subjects	2.20***
		[0.25]
N of decisions		7307
N of participants		234

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; standard errors in brackets; ††random intercept variance within subjects has been fixed to $\pi^2/3=3.29$.

Table S11. Multilevel Mixed-Effects Ordered Logistic Regression – Conditional Effects of Reputation Scores on Gossip Choices; Interaction Effects

Dependent variable: gossip choice (1: ☹, 2: 😐, 3: 😆)	Model 1	Model 2	Model 3
Target's reputation score before update (0-100)	0.05***	0.04***	0.03***
	[0.00]	[0.00]	[0.00]
Manipulation 1 (baseline: low competition)			
high competition	0.78	0.05	0.19
	[0.58]	[0.20]	[0.22]
Manipulation 2 (baseline: control)			
social bonding	-0.02	-0.16	-0.39
	[0.29]	[0.25]	[0.25]
cross-checking	0.38	0.1	-0.32
	[0.38]	[0.27]	[0.25]
Manipulation 1 * Manipulation 2	-0.30		
	[0.27]		
Manipulation 1 * Target's reputation score before update	-0.01*		0.00
	[0.01]		[0.00]
Manipulation 2 * Target's reputation score before update	-0.01***	-0.01***	
	[0.00]	[0.00]	
Manipulation 1 * Manipulation 2 * Target's reputation score before update	0		
	[0.00]		
Rival (0: target is not sender's rival, 1: target is sender's rival)	-0.29***	-0.29***	-0.29***
	[0.06]	[0.06]	[0.06]
Target's PD decision in the same round (baseline: not a PD partner)			
cooperate	1.36***	1.36***	1.34***
	[0.15]	[0.15]	[0.15]
defect	-1.36***	-1.36***	-1.36***
	[0.11]	[0.11]	[0.11]
Round	-0.02***	-0.02***	-0.02***
	[0.01]	[0.01]	[0.01]
cutoff1	-0.71**	-0.91***	-1.03***
	[0.27]	[0.23]	[0.23]
cutoff2	1.30***	1.10***	0.97***
	[0.27]	[0.23]	[0.23]
random intercept variance between subjects	2.20***	2.21***	2.20***
	[0.25]	[0.25]	[0.25]
N of decisions	7307	7307	7307
N of participants	234	234	234

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; standard errors in brackets; ††random intercept variance within subjects has been fixed to $\pi^2/3=3.29$.

Table S12. Multilevel Mixed-Effects Ordered Logistic Regression – Conditional Effects of PD decisions on Gossip Choices; Interaction Effects

Dependent variable: gossip choice (1: 😟, 2: 😐, 3: 😊)	Model 1	Model 2	Model 3
Target’s PD decision in the same round (baseline: not a PD partner)			
cooperate	1.06***	1.26***	1.33***
	[0.24]	[0.19]	[0.16]
defect	-1.96***	-1.56***	-1.38***
	[0.41]	[0.29]	[0.15]
Manipulation 1 (baseline: low competition)			
	0.25	0.06	0.05
	[0.54]	[0.20]	[0.20]
Manipulation 2 (baseline: control)			
social bonding	-0.34	-0.39	-0.39
	[0.28]	[0.25]	[0.25]
cross-checking	-0.24	-0.33	-0.32
	[0.35]	[0.25]	[0.25]
Manipulation 1 * Manipulation 2	-0.1		
	[0.25]		
Manipulation 1 * Target’s PD decision in the same round	0.4		0.03
	[0.28]		[0.10]
Manipulation 2 * Target’s PD decision in the same round	0.14	0.05	
	[0.09]	[0.06]	
Manipulation 1 * Manipulation 2 * Target’s PD decision in the same round	-0.18		
	[0.13]		
Rival (baseline: target is not sender’s rival)			
target is sender’s rival	-0.29***	-0.29***	-0.29***
	[0.06]	[0.06]	[0.06]
Target’s reputation score before update (0-100)	0.03***	0.03***	0.03***
	[0.00]	[0.00]	[0.00]
Round	-0.02***	-0.02***	-0.02***
	[0.00]	[0.00]	[0.00]
cutoff1	-1.05***	-1.10***	-1.09***
	[0.26]	[0.23]	[0.22]
cutoff2	0.95***	0.90***	0.91***
	[0.26]	[0.22]	[0.22]
random intercept variance between subjects	2.20***	2.20***	2.20***
	[0.25]	[0.25]	[0.25]
N of decisions	7307	7307	7307
N of participants	234	234	234

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; standard errors in brackets; ††random intercept variance within subjects has been fixed to $\pi^2/3=3.29$.

Table S13. Multilevel Mixed-Effects Linear Regression – Conditional Effect of Gossip on Reputation Assignment; Interaction Effects

Dependent variable: How much subject trust others (0-100)	Model 1	Model 2	Model 3
Received Gossip (baseline: 😐)†			
no gossip about others	2.45*	4.11***	1.75***
	[0.98]	[0.74]	[0.47]
😊	7.61***	8.43***	7.19***
	[0.69]	[0.61]	[0.53]
☹	-5.23***	-6.08***	-4.88***
	[0.76]	[0.69]	[0.63]
Manipulation 1 (baseline: low competition)			
high competition	0.6	-0.26	-0.17
	[2.19]	[0.82]	[0.83]
Manipulation 2 (baseline: control)			
social bonding	-1.92	-2.12*	-2.04*
	[1.13]	[1.00]	[1.00]
cross-checking	-2.92*	-3.31**	-3.16**
	[1.43]	[1.02]	[1.02]
Manipulation 1 * Manipulation 2	-0.39		
	[1.02]		
Manipulation 1 * Received Gossip	-1.67*		-0.49*
	[0.65]		[0.24]
Manipulation 2 * Received Gossip	0.17	0.48**	
	[0.22]	[0.15]	
Manipulation 1 * Manipulation 2 * Received Gossip	0.61*		
	[0.31]		
Round	-0.14***	-0.14***	-0.14***
	[0.01]	[0.01]	[0.01]
Other Players' Reputation Score in the Previous Round	0.69***	0.69***	0.69***
	[0.00]	[0.00]	[0.00]
PD Decision (baseline: not a PD partner)			
cooperate	8.64***	8.64***	8.63***
	[0.35]	[0.35]	[0.35]
defect	-9.07***	-9.06***	-9.07***
	[0.27]	[0.27]	[0.27]
Rival (baseline: not rival)			
rival	-2.05***	-2.05***	-2.05***
	[0.15]	[0.15]	[0.15]
Gossip Sender			
sent messages	0.58*	0.58*	0.59*
	[0.24]	[0.24]	[0.24]
did not sent any messages	-3.84***	-3.84***	-3.86***
	[0.73]	[0.73]	[0.73]
constant	14.81***	13.38***	15.63***

	[1.38]	[1.12]	[0.98]
random intercept variance between subjects	1.83***	1.83***	1.83***
	[0.05]	[0.05]	[0.05]
random intercept variance within subjects	2.95***	2.95***	2.95***
	[0.00]	[0.00]	[0.00]
N of decisions	79511	79511	79511
N of participants	234	234	234

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, standard errors in brackets; †Mixed gossip about the same target was coded as follows: 😊😐→😊, ☹😐→☹, 😊😊→😐

Table S14. Multilevel Mixed-Effects Linear Regression – Conditional Effect of Gossip Sender's reputation on Reputation Assignment to the Target of Gossip; Interaction Effects

Dependent variable: How much subject trust others (0-100)		Model 1
Target's Reputation Score in the Previous Round		0.61***
		[0.01]
PD Decision (baseline: not a PD partner)		
	cooperate	9.16***
		[1.33]
	defect	-8.73***
		[1.04]
Rival (0:no, 1:yes)		-2.53***
		[0.59]
Received Gossip (baseline: 😐) †		
	😊	6.74***
		[0.77]
	😟	-3.20***
		[0.84]
Round		-0.18***
		[0.05]
Manipulation 1 (0: low competition, 1: high competition)		-1.21
		[1.12]
Manipulation 2 (baseline: control)		
	social bonding	-1.35
		[1.35]
	cross-checking	-2.74*
		[1.39]
Sender's Reputation Score		0.05*
		[0.02]
Sender's Reputation Score – Received Gossip interaction		-0.03**
		[0.01]
	constant	19.03***
		[1.57]
	random intercept variance between subjects	2.02***
		[0.06]
	random intercept variance within subjects	3.10***
		[0.01]
N of decisions		7304
N of participants		234

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, †Mixed gossip about the same target was coded as follows: 😊😐→😊, 😟😐→😟, 😊😊→😐

Table S15. Multilevel mixed-effects logistic regression – Conditional Effect of PD Partner's Reputation on PD decisions; Interaction Effects

Dependent variable: PD choice (0: defect, 1: cooperate)	Model1	Model2	Model3	Model4
Partners Reputation Score	0.02***	0.02***	0.02***	0.03***
	[0.00]	[0.00]	[0.00]	[0.00]
Manipulation 2 (baseline: control)				
social bonding	-0.05	-0.36	-0.18	-0.05
	[0.36]	[0.43]	[0.37]	[0.36]
cross-checking	-0.64	-1.24*	-0.88*	-0.65
	[0.37]	[0.57]	[0.39]	[0.37]
Manipulation 1 (baseline: low competition)				
high competition	0.94**	0.87	0.95**	1.42***
	[0.30]	[0.85]	[0.30]	[0.32]
Manipulation 1 * Manipulation 2		0.29		
		[0.39]		
Manipulation 1 * Partners Reputation Score		0		-0.01***
		[0.01]		[0.00]
Manipulation 2 * Partners Reputation Score		0.00*	0	
		[0.00]	[0.00]	
Manipulation 1 * Manipulation 2 * Partners Reputation Score		0		
		[0.00]		
Relative Reputation Score†	-0.04	-0.04	-0.04	-0.04
	[0.02]	[0.02]	[0.02]	[0.02]
Relative Reputation Score squared	0	0	0	0
	[0.00]	[0.00]	[0.00]	[0.00]
Player Played PD as a Receiver	0.08	0.08	0.08	0.08
	[0.06]	[0.06]	[0.06]	[0.06]
Round	-0.06***	-0.06***	-0.06***	-0.07***
	[0.01]	[0.01]	[0.01]	[0.01]
constant	-1.06*	-1.01*	-0.95*	-1.28**
	[0.43]	[0.48]	[0.44]	[0.44]
random intercept variance between subjects ††	5.17***	5.19***	5.18***	5.15***
	[0.65]	[0.65]	[0.65]	[0.64]
N of decisions	8492	8492	8492	8492
N of participants	234	234	234	234

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, standard errors in brackets; †Original scale contains integers (-24 to +24) and was converted to natural numbers (1 to 48); ††random intercept variance within subjects has been fixed to $\pi^2/3$=3.29.

Table S16. Multilevel mixed-effects logistic regression – Conditional Effect of Round on PD decisions; Interaction Effects

Dependent variable: PD choice (0: defect, 1: cooperate)	Model 1	Model 2	Model 3
Round	-0.03	-0.04**	-0.11***
	[0.02]	[0.02]	[0.01]
Manipulation 1 (baseline: low competition)			
high competition	0.82	0.90**	0.06
	[0.89]	[0.29]	[0.33]
Manipulation 2 (baseline: control)			
social bonding	0.26	0.03	-0.2
	[0.43]	[0.37]	[0.36]
cross-checking	0.05	-0.38	-0.86*
	[0.58]	[0.41]	[0.36]
Manipulation 1 * Manipulation 2	-0.38		
	[0.41]		
Manipulation 1 * Round	-0.02		0.06***
	[0.03]		[0.01]
Manipulation 2 * Round	-0.04***	-0.02*	
	[0.01]	[0.01]	
Manipulation 1 * Manipulation 2 * Round	0.04**		
	[0.01]		
Player Played PD as a Receiver	0.1	0.1	0.1
	[0.06]	[0.06]	[0.06]
constant	-0.17	-0.39	0.28
	[0.39]	[0.32]	[0.32]
random intercept variance between subjects	5.01***	4.98***	4.99***
	[0.63]	[0.62]	[0.63]
N of decisions	8404	8404	8404
N of participants	234	234	234

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, standard errors in brackets; † random intercept variance within subjects has been fixed to $\pi^2/3=3.29$.

Table S17. Multilevel mixed-effects logistic regression – Conditional Effect of PD Partner's Previous Cooperativeness on PD decisions; Interaction Effects

Dependent variable: PD choice (0: defect, 1: cooperate)	Model 1	Model 2	Model 3	Model 4
PD Partner's Previous Cooperation Rate	1.25***	0.09	0.69	1.22***
	[0.17]	[0.69]	[0.47]	[0.24]
Manipulation 1 (baseline: low competition)				
high competition	0.79**	0.06	0.79**	0.76*
	[0.28]	[0.86]	[0.28]	[0.32]
Manipulation 2 (baseline: control)				
social bonding	-0.24	-0.52	-0.37	-0.24
	[0.34]	[0.41]	[0.36]	[0.34]
cross-checking	-0.71*	-1.26*	-0.94*	-0.71*
	[0.35]	[0.54]	[0.39]	[0.35]
Manipulation 1 * Manipulation 2		0.33		
		[0.39]		
Manipulation 1 * PD Partner's Previous Cooperation Rate		1.15		0.06
		[0.95]		[0.34]
Manipulation 2 * PD Partner's Previous Cooperation Rate		0.54	0.27	
		[0.31]	[0.21]	
Manipulation 1 * Manipulation 2 * PD Partner's Previous Cooperation Rate		-0.52		
		[0.43]		
Player Played PD as a Receiver	0.1	0.1	0.1	0.1
	[0.06]	[0.06]	[0.06]	[0.06]
constant	-1.71***	-1.38***	-1.57***	-1.69***
	[0.30]	[0.38]	[0.31]	[0.30]
random intercept variance between subjects	4.58***	4.58***	4.58***	4.57***
	[0.58]	[0.58]	[0.58]	[0.58]
N of decisions	8404	8404	8404	8404
N of participants	234	234	234	234

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, standard errors in brackets; † random intercept variance within subjects has been fixed to $\pi^2/3=3.29$.

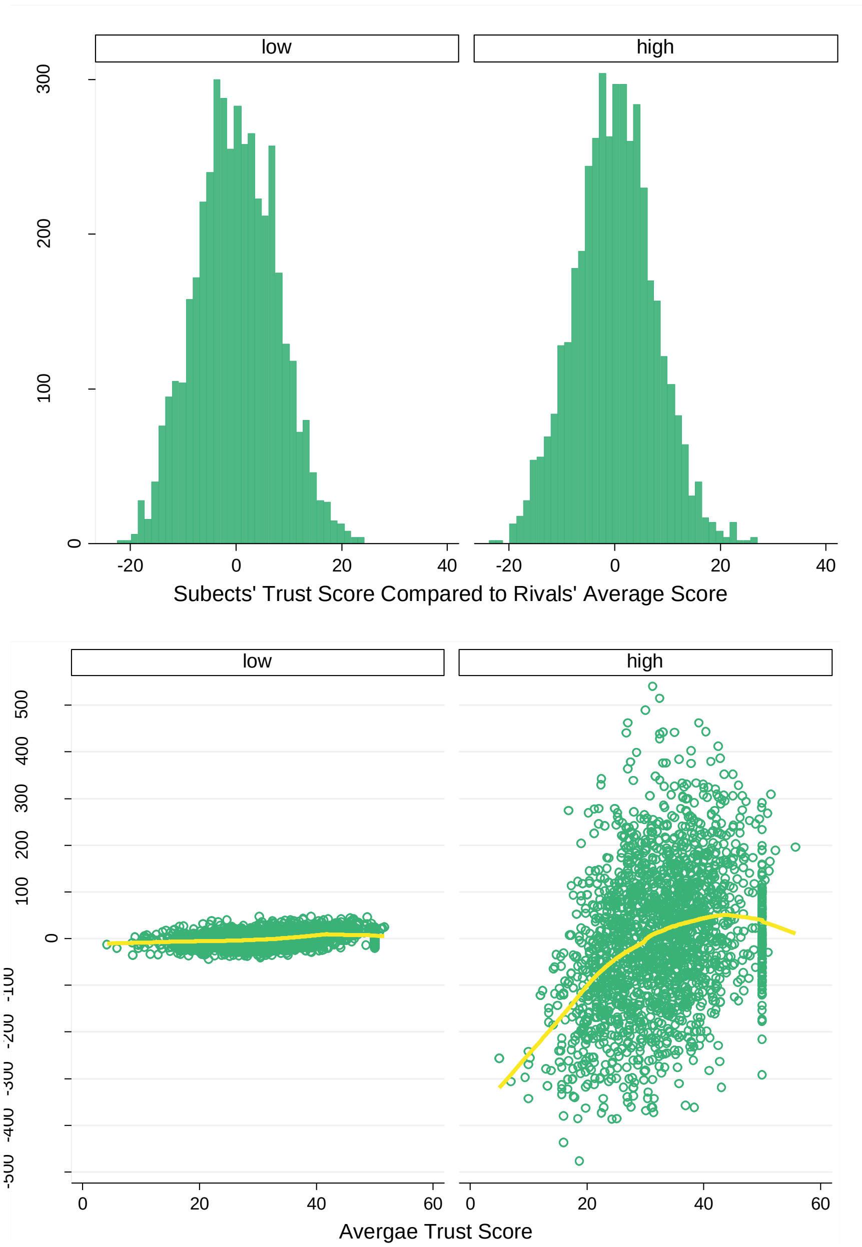


Figure S1 The Distribution of Relative Reputations (a) and its Financial Consequence (b). The distribution of the relative reputations formed was the same in the two groups of Manipulation 1 (right: low competition, left: high competition). Due to Manipulation 1, however, the change in payoffs according to participants' relative reputations were greater.

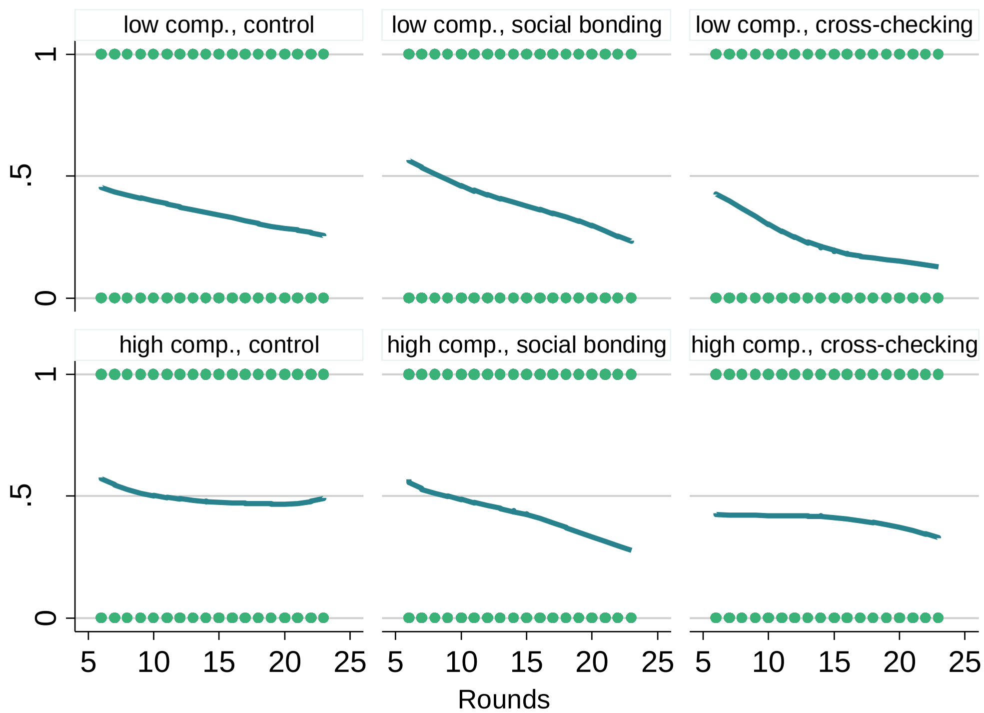
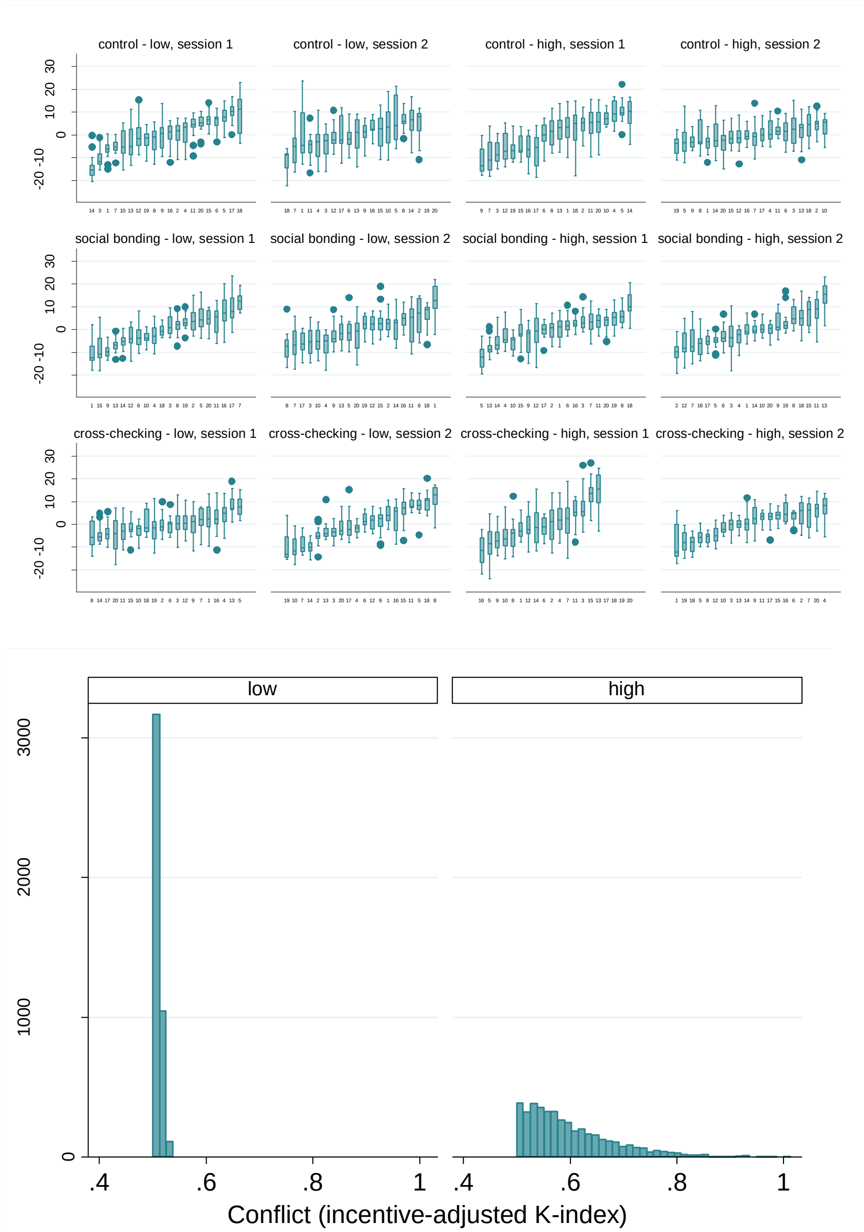


Figure S2 Cooperation by Rounds Using Local Linear Smoothing. Cooperation in high competition was more resistant and remained stable where information flow was high (especially in high competition – control condition).



FigureS3 The Variability of Relative Reputations by Participants in Experimental Sessions (a) and its Effect on Conflict in the PD Game (b). Since relative reputations have not fluctuated much (left), participants could expect stability in their relative position and rewards or punishments individually differentiated the degree of conflict between prosocial and pro-self behaviour (calculated by the incentive-adjusted K-index, Rapoport 1967) (right).

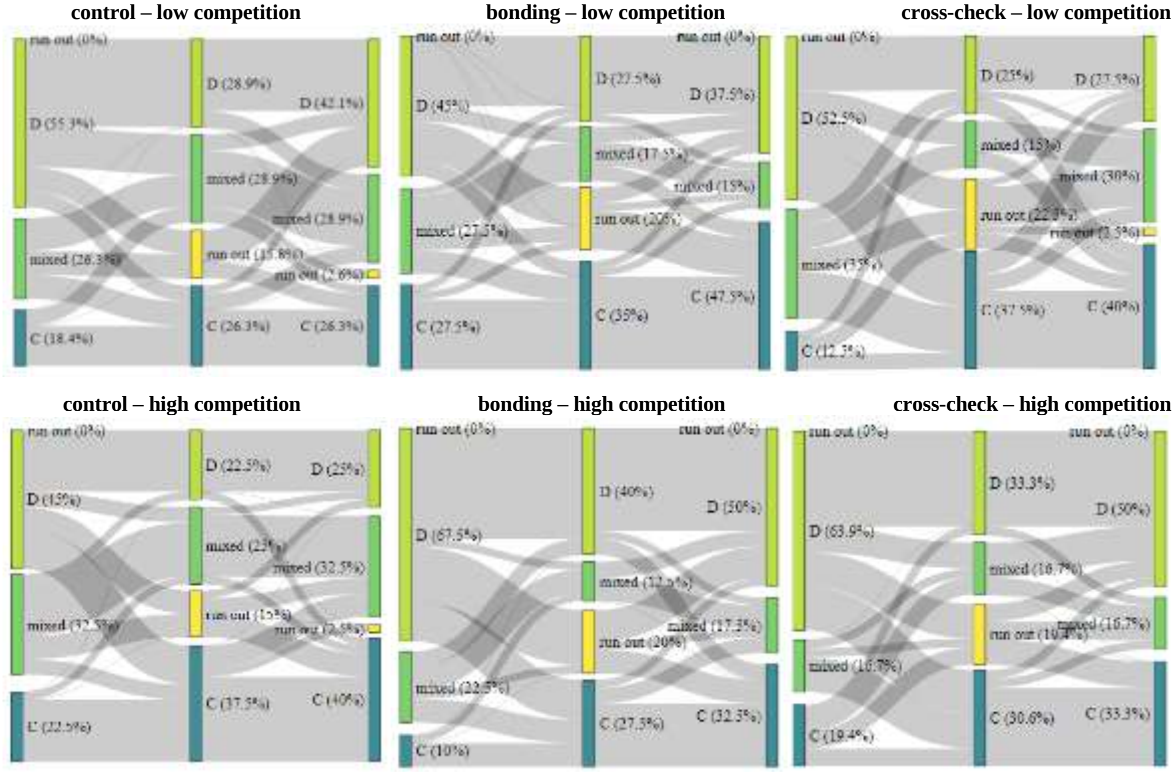


Figure S4 Shifts in PD decisions when Gossip is Introduced. Each stacked bar shows how participants decided in their PD games without communication (Round 5), after they were informed about gossip (Round 6) and after they de facto gossiped (Round 7), respectively. In each round, participants played 2 PDs, therefore D stands for defection in both games, C stands for cooperation and mixed shows when participants played a mixed strategy. Flows represents how participants changed their decisions between these rounds. Although deviation from self-interest become wider with the possibility of gossip exchange (in round 6), after playing the game with gossip and reputation once (Round 7), many people reverted to defection.

S2 Instructions

Welcome to the decision-making experiments organized by the "Lendület" Research Center for Educational and Network Studies (RECENS)!

The decision-making experiments carried out by the staff of the Hungarian Academy of Sciences, Centre for Social Sciences, "Lendület" Research Center for Educational and Network Studies (RECENS): Károly Takács (Principal Investigator, email: karoly.takacs@uni-corvinus.hu), Szabolcs Számadó (senior research fellow, email: szamszab@ludens.elte.hu), and Flóra Samu (PhD student, email: samufloraa@gmail.com). The research is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC_CoG_2014_648693, Principal Investigator: Károly Takács).
The aim of the scientific research is to examine the nature and the conditions of informal communication and empirically test theoretical hypotheses under controlled conditions in decision-making experiments. You can obtain more information on the research and about the researchers involved at the recens.tk.mta.hu website.

In the following, instructions will appear on your screen about the experiment. You participate in the experiment together with the other people in this room. Instructions are the same for all participants. It is important that all information you receive is real. You get the most important instructions on paper as well. You can use them at any time during the experiment. Please do not take these instructions with you after the experiment, leave them on the table.

Please turn off your phone or put it on silent mode! VERY IMPORTANT rule is that it is STRICTLY FORBIDDEN to talk to or to signal to others! Violation of this rule may result in disqualification from this experiment. The experiment lasts for approx. 75 minutes. Your payoffs in the experiment will be paid through a student organization by the 10th day of the month following the experiment. You must make a contract with a student organization. At the end of the experiment, participants will get a certificate about the amount they earned in the experiment. This certificate will not contain the name of the participants. You must bring this certificate to the Student Committee or give it to their representative on the spot.

The amount of your payoff depends on your own choices and the decisions of others. The precise calculation of your payment will be described later in more detail. Payoffs made in the experiment will be shown as gross amounts. You will receive a 1000 HUF show-up fee, which you will receive regardless of your payoffs in the rounds of the experiment. This is also paid through the student co-operative.

During the experiment, we do not handle personal data, the participants participate anonymously in the experiment. At the beginning of the experiment participants draws a random number, by which they are randomly and anonymously assigned to the experimental devices (computer). The leader of the experiment will be able to determine the payoffs by these random numbers. At the end of the experiment, participants will get a certificate about the amount they earned in the experiment. The certificate will be issued by the leader of the experiment. This certificate will not contain the name of the participants, so please keep it safe! Since we do not handle personal data using this method, after the experiment we will not be able to identify your payoff in the experimental database. The data generated by your decisions are therefore not personal.
Data recorded during the experiment do not contain information that can be traced back to individuals. If a participant withdraws participating in the experiment (this can be done at any time during the experiment), then their responses will be deleted according to the IP address of the computer used by the participant. If a participant withdraws his/her participation, payment is not possible.

By signing this consent, you agree that your answers will be exclusively and anonymously used for scientific research purposes.

Thank you in advance for your participation! Have fun and good luck!

INSTRUCTIONS

Participants

The experiment consists of several rounds of decision-making situations. In each round, you will be paired with two other participants. Participants of the experiment are identified with fictitious IDs.
It is important that you do not play with the same participants in each round. Both of your pairs are determined by a random number generator. The IDs of your current pairs will be displayed on your screen. The IDs will be kept confidential: we will not reveal which identifier belongs to which participant neither during nor after the experiment.

The game

With both of your pairs, you face the following decision. You can choose from two options: these two options are indicated with L and R. The amount that you earn in this round does not depend only on your decision, but on the decision of your pair as well. The payment you can expect is:
If both of you select L: HUF 1500
If you select L and your partner selects R: HUF 2000
If you select R and your partner selects L: HUF 0
If both of you select R: HUF 500
If you run out of time: HUF 0

Your final payoff

Each round is important for your final payoff. A random number generator will select 1 round in the first phase and 5 rounds in the second phase. The average of your payoffs in these six rounds will be your final payoff. In addition, we gave everyone 1000 forints as a show-up fee. Payoffs will be rounded up to HUF 100.
It is important that during the experiment you receive real information. In the upper right corner of your screen you can see how much time you have left to make your decision.

If you are ready, please click on the "Next" button on your screen.

Have fun and good luck!

FURTHER INSTRUCTIONS

You must make the same decision in the following rounds and the amounts you can win are the same as previously. In addition to this, the following changes are in place:

From now on you'll have the opportunity to send a message about others to a randomly selected participant. In each round, we randomly pair you with a player to whom you can send the first message, and we also use a random number generator to determine who the message you are sending will be about. On your screen you will see the ID of the participant you can send the message TO and the IDs of the two players you can send the message ABOUT. [*Manipulation 2*]

- *[without manipulations]: After sending the first message, you will have the opportunity to send another message to a randomly selected participant, again about a randomly selected participant.*
- *[bonding]: After sending the first message, you will have the opportunity to send another message to the participant from whom you received the first message, again about a randomly selected participant.*
- *[cross-check] After sending the first message, you will have the opportunity to decide whether you ask more information - from a randomly selected participant - about the player you received the message about.*

There is no cost for sending a message. The time available for your choices will be displayed in the top right corner of your screen. Be careful, if you run out of time, it is considered that you did not want to send a message! After all messages have been sent, the messages you received will appear on your screen.

Apart from the messages, it is also new that in each round you can rate other participants in a 100-point scale depending on how trustworthy you think they are. A maximum of 1900 points can be distributed. By default, the neutral value of 50 is set to everyone.

Similarly, others can evaluate YOUR trustworthiness. The points you receive will be taken into account in your final payoff! Based on your mean score, we will adjust your earning in each round as follows: Your winnings will depend on the scores of five other players who will be randomly selected for you at the very beginning of this phase.

These 5 players will remain the same until the end of the experiment. If you get the same average score from others as these five players get on average, your earning in that round will not change. A one-unit decrease / increase in your mean score compared to theirs decreases / increases your payment by HUF *[Manipulation 1]*.

- *[low-intensity] 2*
- *[high-intensity] 20*

For instance, if the other participants gave you an average of 0 and the other 5 players got an average of 100, then your payment will be reduced by HUF [200 /2000] (see Example 1 in Table 1). On the contrary if other participants gave you an average of 100 and the other five players got an average of 0, then your payment will increase by HUF [200 /2000] (see Example 3 in Table 1). It is important that the points you gave to the five players do not considered in their average score.

Table 1 Possible changes in your payment

	Your average score	**Total score of the five players**	Difference	**Change in your payment**
Example 1	0	100	-100	-[200/2000] Ft
Example 2	50	50	0	0 Ft
Example 3	100	0	+100	+[200/2000] Ft

The second part of the experiment also consists of several decision-making rounds. As before, we will pair you with two other participants in each round. Pairing this time, however, will be based on encounters in the previous round. Half of the participants will play with the player who sent a message and with who a message was about. Therefore, the other half will be paired with the person to whom they have sent a message and who have received a message about them. In each round a random number generator decides which half of the participants you will belong to.

If you are ready press the "Next" button! Have fun and good luck!

S3 Screens of one round in the first phase (Round 1-5) and second phase (Round 6 -23) of the experiment

Round 1	Remaining time (sec): 19

This round may count towards your final payment. Remember, your payoff is:
If both of you choose L: 1500
If you choose L and your partner chooses R: 0
If you choose R and your partner chooses L: 2000
If both of you choose R: 500

Your ID: Amalthea

Your first Partner's ID: **Daphnis**
Which option do you choose against her/him? L/R

Your second Partner's ID: **Metis**
Which option do you choose against her/him? L/R

To confirm your decision, you need to click on the 'confirm' button.

Confirm

Your decision against your first partner in this round: R
As a reminder, your partner's ID: **Daphnis**
Your partner's decision in this round: R
Your payoff in this round: 500

Your decision against your second partner in this round: L
As a reminder, your partner's ID: **Metis**
Your partner's decision in this round: R
Your payoff in this round: 0

Your combined payoff in this round (in HUF): 250

OK

Round 10

Remaining time (sec): 8

This round may count towards your final payment.
Remember, your payoff is:
If both of you choose L: 1500
If you choose L and your partner chooses R: 0
If you choose R and your partner chooses L: 2000
If both of you choose R: 500

Your ID: Amalthea

Your first Partner's ID: **Kale**
Which option do you choose against her/him? L/R

Your second Partner's ID: **Galatea**
Which option do you choose against her/him? L/R

To confirm your decision, you need to click on the 'confirm' button.

Confirm

ID	Scores
Bestla	0
Chaldene	50
Daphnis	48
Elara	50
Farbauti	50
Galatea	50
Halimede	50
Iocaste	50
Jarnsaxa	50
Kale	100
Lysithea	50
Metis	0
Nereid	50
Orthosie	50
Pasiphae	48
Rhea	50
Sinope	0
Thyone	50
Umbriel	50

Your decision against your first partner in this round: L
As a reminder, your partner's ID: Kale
Your partner's decision in this round: R
Your payoff in this round: 0

Your decision against your second partner in this round: L
As a reminder, your partner's ID: Galatea
Your partner's decision in this round: R
Your payoff in this round: 0

Your combined payoff in this round (in HUF): 0

OK

Round 10		Remaining time (sec): 8	
As a reminder, your partner's ID: Amalthea Your first Partner's ID: **Kale** Your partner's decision in this round: R Your second Partner's ID: **Galatea** Your partner's decision in this round: R	Now you can send a message to a randomly selected participant. If your time expires, we consider it as you did not want to send a message. Receiver: **Iocaste** The message is about: **Nereid** What message would you send? ○ :) ○ :| ○ :(Next	**ID**	**Scores**
We calculate your payoff relative to the scores of the following participants: Bestla Elara Halimede Jarnsaxa Rhea		Bestla	0
		Chaldene	50
		Daphnis	48
		Elara	50
		Farbauti	50
		Galatea	50
		Halimede	50
		Iocaste	50
		Jarnsaxa	50
		Kale	100
		Lysithea	50
		Metis	0
		Nereid	50
		Orthosie	50
		Pasiphae	48
		Rhea	50
		Sinope	0
		Thyone	50
		Umbriel	50

<table>
<tr><td>Round 10</td><td></td><td colspan="2">Remaining time (sec): 8</td></tr>
<tr><td>As a reminder, your partner's ID: Amalthea
Your first Partner's ID: Kale
Your partner's decision in this round: R
Your second Partner's ID: Galatea
Your partner's decision in this round: R</td>
<td rowspan="3">[Control treatment:] Now you can send another message to a randomly selected participant. You can send the message about another randomly selected participant.
Receiver: Umbriel
The message is about: Metis
What message would you send?
○ :)
○ :|
○ :(
[Bonding manipulation:] Now you can send a message to the sender of your previous message.
Receiver: Chaldene
The message is about: Pasiphae
What message would you send?
○ :)
○ :|
○ :(
[Cross-check manipulation:] Now you can request additional information about the participant you received the previous message about. You can request information from a randomly selected participant.
From whom you can request information:
Bestla
Who you can ask for information about:
Rhea
Do you want more information?
○ Yes
○ No
Next</td>
<td rowspan="3">ID
Bestla
Chaldene
Daphnis
Elara
Farbauti
Galatea
Halimede
Iocaste
Jarnsaxa
Kale
Lysithea
Metis
Nereid
Orthosie
Pasiphae
Rhea
Sinope
Thyone
Umbriel</td>
<td rowspan="3">Scores
0
50
48
50
50
50
50
50
50
100
50
0
50
50
48
50
0
50
50</td></tr>
<tr><td>We calculate your payoff relative to the scores of the following participants:
Bestla
Elara
Halimede
Jarnsaxa
Rhea</td></tr>
<tr><td>You received a message FROM:
Chaldene
About: Rhea
The content of the message:
:(</td></tr>
</table>

Note: Even if subjects didn’t get messages, they had the option to send a second message, except in the cross-check treatment, it was only available where upon request.

<table>
<tr><td>Round 10</td><td></td><td colspan="2">Remaining time (sec): 8</td></tr>
<tr><td>As a reminder, your partner's ID:
Amalthea
Your first Partner's ID: Kale
Your partner's decision in this round: R
Your second Partner's ID: Galatea
Your partner's decision in this round: R</td><td rowspan="3">[Cross-check manipulation:]
Now you can send a message to the participant who requested information from you.
Receiver: Halimede
The message is about: Jarnsaxa
What message would you send?
○ :)
○ :|
○ :(
Next</td><td rowspan="3">ID
Bestla
Chaldene
Daphnis
Elara
Farbauti
Galatea
Halimede
Iocaste
Jarnsaxa
Kale
Lysithea
Metis
Nereid
Orthosie
Pasiphae
Rhea
Sinope
Thyone
Umbriel</td><td rowspan="3">Scores
0
50
48
50
50
50
50
50
50
100
50
0
50
50
48
50
0
50
50</td></tr>
<tr><td>We calculate your payoff relative to the scores of the following participants:
Bestla
Elara
Halimede
Jarnsaxa
Rhea</td></tr>
<tr><td>You received a message FROM:
Chaldene
About: Rhea
The content of the message:
:(

You received a message FROM:
Halimede
About: Jarnsaxa
The content of the message:
Request for information</td></tr>
</table>

Note: only in cross-check treatment

Round 10

Remaining time (sec): 8

As a reminder, your partner's ID: Amalthea
Your first Partner's ID: **Kale**
Your partner's decision in this round: R
Your second Partner's ID: **Galatea**
Your partner's decision in this round: R

We calculate your payoff relative to the scores of the following participants:
Bestla
Elara
Halimede
Jarnsaxa
Rhea
It is important that the scores you gave to these five players do not considered in their average score.

You received a message FROM: **Chaldene**
About: **Rhea**
The content of the message:
:(

You received a message FROM: **Bestla**
About: **Rhea**
The content of the message:
:|

How much do you trust other participants? In the middle of the screen you can now change their scores.

Bestla	___
Chaldene	___
Daphnis	___
Elara	___
Farbauti	___
Galatea	___
Halimede	___
Iocaste	___
Jarnsaxa	___
Kale	___
Lysithea	___
Metis	___
Nereid	___
Orthosie	___
Pasiphae	___
Rhea	___
Sinope	___

Next

ID	**Scores**
Bestla	0
Chaldene	50
Daphnis	48
Elara	50
Farbauti	50
Galatea	50
Halimede	50
Iocaste	50
Jarnsaxa	50
Kale	100
Lysithea	50
Metis	0
Nereid	50
Orthosie	50
Pasiphae	48
Rhea	50
Sinope	0
Thyone	50
Umbriel	50

Your average score: 50.0
The average of those participants against whom we calculate your payoff:
The average of Bestla: 50.0
The average of Elara: 50.0
The average of Halimede: 50.0
The average of Jarnsaxa: 35.7
The average of Rhea: 25.0

In total, their average score: 42.1

Your winnings (in HUF) in this round: 158

OK

4. CONDITION-DEPENDENT TRADE-OFFS MAINTAIN HONEST SIGNALING[10]

4.1. Introduction

To understand the evolution of communication in animals and humans, it is essential to find out why signals of senders should be trusted by receivers. Honest signalling is challenging to explain if some senders, for instance in the absence of a relevant quality, have an interest to deceive the receivers. The receivers face the problem of how to differentiate between senders with and without the relevant quality if both send the same signal. Yet signals are plentiful in both nature and in human interactions and many signals are informative to the receiver. For example, in birds of paradise, males display complex dance moves to attract their mate. Why did such elaborated signals evolve and what do they signal? Why should a female pick a male with an extraordinary display?

Claiming a solution to the puzzle, the Handicap Principle proposed that to be honest, signals need to be costly to produce, and consequently they have to function as handicaps that only high-quality individuals can bear (Zahavi, 1975; Grafen, 1990; Zahavi and Zahavi, 1997). For instance, the peacock's train can serve as an honest signal of peacock's quality because it is costly to produce (Zahavi, 1975).

Zahavi's idea remained contested (Maynard Smith, 1976; Kirkpatrick, 1986) until Grafen (1990) published an analytical model in which he claimed to show that honest biological signals have to function as handicaps ("if we see a character which does signal quality, then it must be a handicap" Grafen 1990; page 521). The Handicap Principle remained highly influential in biology despite criticism through the years (Maynard Smith, 1976; Kirkpatrick, 1986; Maynard Smith and Harper, 2003; Searcy and Nowicki 2005; Biernaskie et al., 2018; Lachmann et al., 2001; Getty, 2006;

---

[10] Számadó, S., Samu, F., & Takács, K. (2022). Condition-dependent trade-offs maintain honest signalling. Royal Society Open Science, 9(10), 220335.

Számadó, 2003,2011; Higham, 2014; Számadó and Penn, 2015, 2018; Penn and Számadó, 2020, Barker et al., 2019; Stibbard-Hawkes, 2019) as researchers still interpret their results in light of the Handicap Principle (e.g. Clifton et al., 2016; Polnaszek and Stephens, 2014; Beecher, 2021). Zahavi's idea was adopted in other disciplines as well, such as anthropology (Bliege Bird, Smith and Bird, 2001; Hawkes and Bliege Bird, 2002; McAndrew, 2019). It inspired studies on human mate choice (Geary et al., 2004; Gangestad and Scheyd, 2005; Frederick and Haselton, 2007), generosity (Boone, 1998; Roberts, 1998), consumer behaviour (Griskevicius et al., 2010), big game hunting (Bliege Bird, Smith and Bird, 2001; Hawkes and Bliege Bird, 2002), trophy hunting (Darimont et al., 2017), risk taking in young man (Nell, 2002; Farthing, 2005), criminal behaviour (Densley, 2012), citizen behaviour (Salamon and Deutsch, 2006); blood donations (Lyle et al., 2009), over-consumption of resources (Sivanathan and Pettit, 2010) and religious rituals (Sosis, 2003; Sosis and Bressler, 2003; Kantner and Vaughn, 2012) just to name a few examples (see McAndrew, 2019 for review).

Grafen's claim, however, was not substantiated as Grafen's model is not a handicap model at all and Grafen's main handicap results are unsupported by his own model (see Penn and Számadó, 2020 for more detail). On the one hand, Grafen (1990) correctly identified the need of differential marginal costs as a mechanism that can maintain honest signalling. On the other hand, (i) this insight was not novel as it is an insight learnt also by the theory of signalling in economics (see Spence, 1973; Riley, 1979). Spence, however, has never proposed that individuals must pay a handicap cost in order for the signal to be honest (e.g., Stibbard-Hawkes, 2019). Subsequent game theoretical models have made it clear that it is not the equilibrium cost of the signal that maintains the honesty of communication but the difference of marginal costs to marginal benefits by signaller type (Hurd, 1995; Számadó, 1999; Lachmann et al., 2001; Bergstrom et al., 2002). That is, honesty can only be guaranteed if the marginal benefits of giving the signal are higher than the marginal costs of transmission for individuals with the inherent quality, while the opposite holds for individuals without the quality. This means that honest signals need not be costly to produce at the equilibrium not even under conflict of interest (Számadó, 1999; Lachmann et al., 2001; Számadó, 2011; Higham, 2014). Highly costly signals could be sent by signallers with the relevant quality and these signals, such as earning a university diploma, will be trusted by receivers because the difference in the benefits and production costs for individuals with the relative quality is higher than for individuals without the quality. But cost-free or even negative cost signals can be honest and evolutionarily stable as well (Enquist, 1985; Hurd, 1995; Számadó, 1999 Lachmann et al., 2001; Számadó, 2011; Számadó et al., 2019). Telling about a good grade earned at school to the parents is cost-free or even self-rewarding and can expected to be believed because the marginal benefits to marginal costs are higher for those who actually earned a good grade compared to those who did not. Cost-free and self-rewarding signals are not rare in nature, for instance, singing to attract mates or enjoying the sun on highest panoramic cliff to signal dominance.

Note, that most of signalling games in economics have taken a different route (compared to biology) and investigated the relevance of 'cheap talk' models. Cheap talk is defined as a message where "players' messages have no direct payoff implications (Crawford, 1998). In a seminal model of the field Crawford and Sobel (1982) were able to show that such cheap talk can be informative as long as the preferences of the signaller and receiver are similar. This inspired a long line of research asking what kind of equilibria are possible in cheap talk games and out of these what can be observed in humans (Dickhaut et al., 1995; Blume et al., 1998; Gneezy, 2005; Cai et al., 2006; Sanchez-Pages and Vorsatz, 2007; see Blume et al., 2020 for a review). Our approach differs from this line of experiments by investigating the role of signal costs and trade-offs.

The closest to our approach is a recent experiment by Rubin et al. (2020) in which they investigated the possibility of hybrid equilibria arising in which signals are not fully but partially informative in a signal-response game with conflict interest in a lab setting. They found that partially informative communication has developed. As a main difference to our investigation, in their experiment they contrasted the hybrid equilibrium manipulation with a simple control condition in which interests were aligned to allow fully informative communication. Hence, they did not systematically test if (partially) honest signalling emerged due to costly signalling or due to signalling trade-offs. All in all, experimental evidence is still lacking whether signals used by humans need to be costly to produce to be honest, or not. To fill this gap, we conducted a laboratory experiment where human participants played a signalling game.

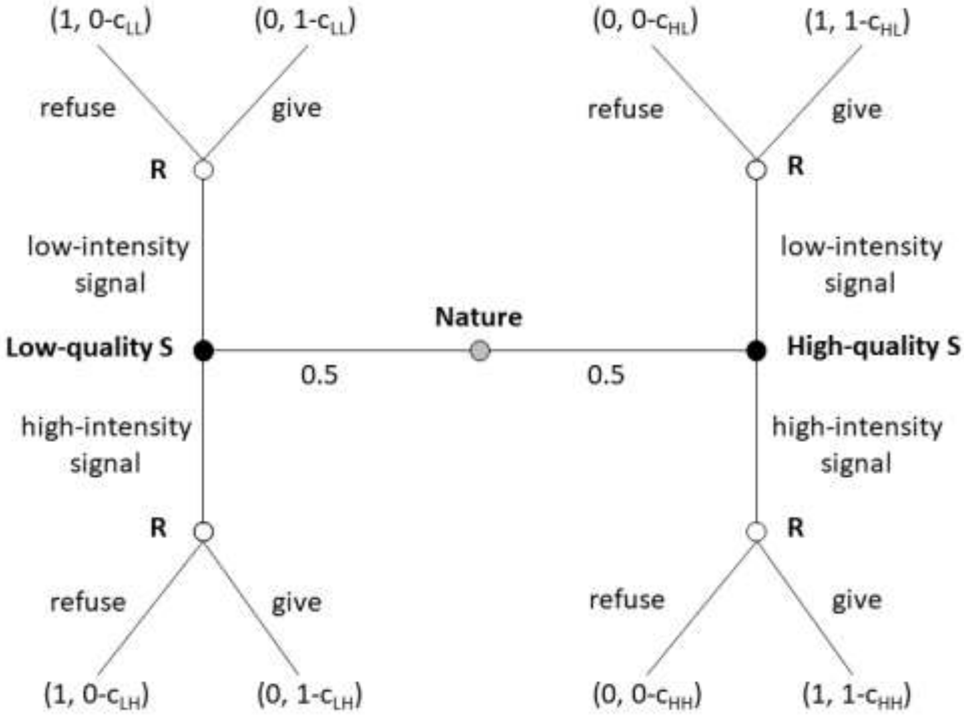


Figure 1 The signalling game with conflict of interest. A differential cost model when the cost is condition-dependent. After nature randomly divides the roles (S and R) and the types of signallers (high- and low-quality S), players make sequential decisions. First, S can send two different signals (low-intensity signal, high-intensity signal) to get the resource from R. After this, R receives the signal and decides whether to provide the resource to S or not. At the end of each round players get feedback on their success. Both high- and low-quality S wins when gets the resource from R, but for R only the dedication of the resource to a high-quality S and its protection from low-quality S generate successful outcomes (see outcomes in brackets at the ends of the decision tree where the first element refers to R and the second to S). The reward for success is indicated simply as 1 in the Figure, in the experiment it meant a payoff of HUF 1200. In the figure cLL, cLH, cHL, and cHH indicate the different cost of signals according to the type of signaller.

4.1.1. Signalling games and terminology

In general, signalling games have been fruitful to model various dyadic situations of communication and to provide insights about the conditions of honesty and reliability of messages. The signal-response game describes the interaction of a signaller (S) and a receiver (R) under the assumption of asymmetrical information, namely that the signaller knows its own quality relevant to the receiver (high vs. low quality), while the receiver does not (see Figure 1). The signaller may or may not give a signal (high vs. low-intensity signal), finally the receiver may or may not transfer the resource to the signaller. The hidden type of the signaller is important here since R gets the highest payoff if it gives the resource to a high-quality S instead of a low-quality one. At the same time, however, both signallers are interested in receiving the resource from R (there is a conflict of interest). Since R does not have information on the type of S, R is trying to predict the type of the signaller from the signal while a low-quality S is trying to conceal it, because otherwise, with ‘honest’ communication, they would not receive the resource from R.

The focus of signalling games is to investigate the equilibrium conditions of honest communication where the receiver can tell the unseen quality of the signaller from the use of (different) signals, i.e., there is a correlation between quality and signal intensity. At the honest equilibrium, the signal will be informative, because the receiver can infer from the intensity of the signal the quality of the sender. Hence, the receiver will trust the signal, i.e., R will transfer the resource to signaller using a high-intensity signal and R will reject signallers using a low-intensity signal. Thus, as a result of honesty and trust, the receiver will be able to accurately transfer the resource to high-quality signallers (accuracy). An equilibrium strategy pair is defined as a set of strategies for the signaller and for the receiver such that it does not benefit to deviate unilaterally for any of the players. The equilibrium cost of signals is the signal cost observed at equilibrium. Since at the honest equilibrium only high-quality S will use the high-intensity signal, the observed equilibrium cost is the cost of high-intensity signals for high-quality type signallers. It is natural to believe that in any signalling system it takes time to realize the equilibrium strategies and hence signal-response games are typically played repeatedly in the lab setting.

The honest equilibrium is often called as "separating" as signaller types are separated by their use of signal (Bergstrom et al., 2002), which leads to a consequence that signals identify the type of the signaller. There can be also other type of equilibria such as "pooling", where some or all of the signaller types are using the same signal (Lachmann and Bergstrom, 1997) or "hybrid", where some types can be mixing different signals with a given probability (Zollman et al., 2013). Note that these later equilibria (pooling and hybrid) will not be entirely honest, however, they need not be (and probably will not be) entirely uninformative either.

Last but not least Figure 2 describes the relation between various costs and the concept of trade-off. Equilibrium cost describes the cost paid by High quality signallers for the use of High-intensity signal (it describes a state). Marginal cost describes the (cost) difference between switching from one type of signal to the other (it describes change in one function). Finally, a trade-off describes a specific relation of two functions: increasing the value of one function is not possible without decreasing the other.

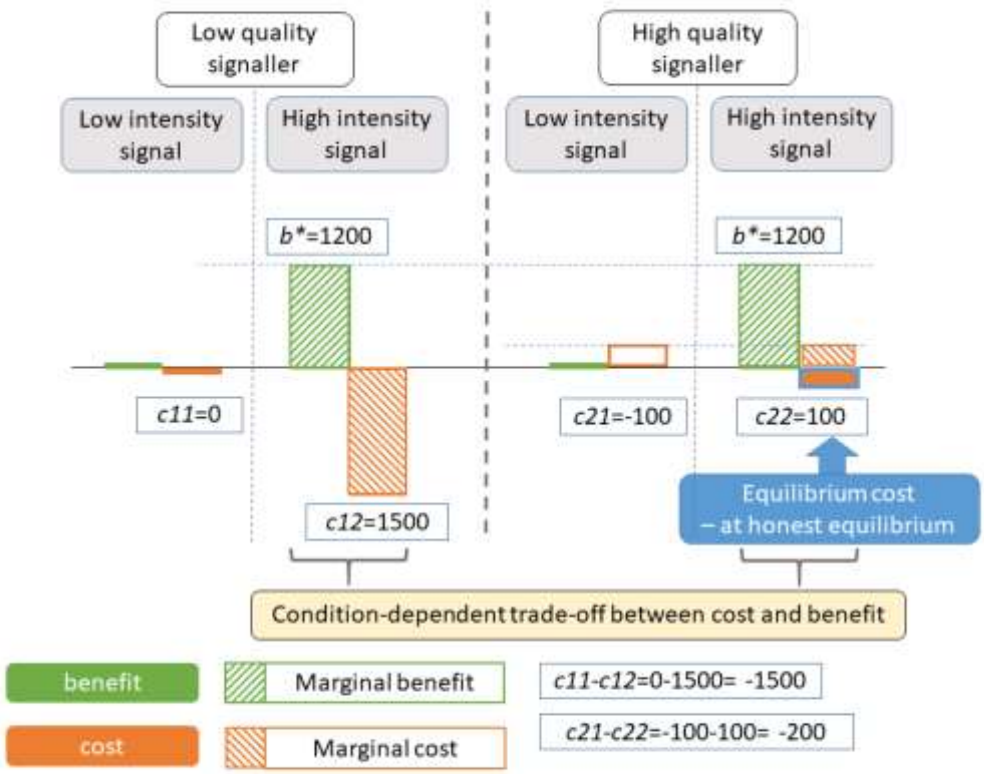


Figure 2. The relation between various costs and the concept of conditional trade-off. Equilibrium cost

describes the cost paid by high-quality signallers for the use of high-intensity signal (it describes a state). Marginal cost describes the (cost) difference between switching from one type of signal to the other (it describes change in one function). Trade-off describes the relation of two functions: increasing the value of one function is not possible without decreasing the other.

4.2. Manipulations and hypotheses

In our experiments, by using a two-factorial experimental design we varied both the production cost of the signal for high-quality signallers and the presence of signalling trade-offs (see Table 1). Signallers could either have high- or low-quality. First, in the signalling cost manipulation (between rows in Figure 3.A) we varied whether (i) high-quality signallers had to pay to use the high-intensity signal; (ii) high-quality signallers could use the high-intensity signal for free; (iii) or high-quality signallers received a benefit (payment) for using the high-intensity signal. Second, in the trade-off manipulation (between columns in Figure 2.A), we tested whether the absence, the presence, or the differentiation of difference between the costs of the two signals by the quality of signallers influence the emergence of honest communication. Accordingly, we had three trade-off conditions: (i) no trade-off, where the cost of low- and high-intensity signal was the same regardless of the quality of the signaller; (ii) trade-off condition, where sending the high-intensity signal was costlier for both low- and high-quality signallers in the same way (i.e. the marginal cost of sending a high-intensity signal was the same for all signallers); (iii) condition-dependent trade-off condition, where the cost of sending a high-intensity signal was relatively costlier for low-quality signallers thank for high-quality signallers (i.e., the marginal cost of sending a high-intensity signal and the cost of sending the low-intensity signal was higher for low-quality signallers than for high-quality signallers). Note that taking the high-intensity signal is never cheaper than taking the low-intensity signal. As a consequence, both low- and high-quality signallers would be motivated to send the low-intensity signal.

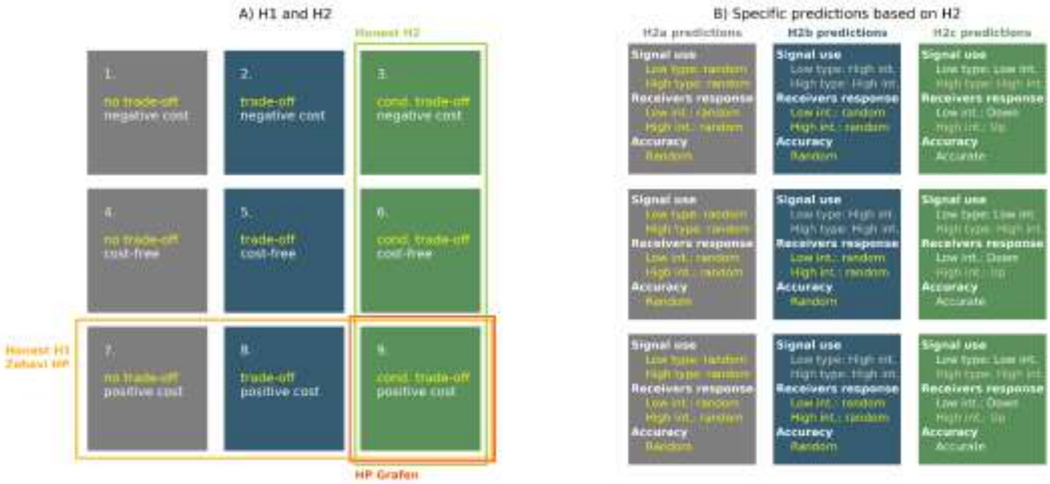


Figure 3. General and specific predictions (H1 and H2). A) The design matrix with 9 manipulations. Rows: equilibrium signals with negative cost, cost-free, positive cost. Columns: signals without trade-off, with trade-off and with conditional trade-off. H1: prediction of Zahavi's Handicap Principle; honesty is maintained by positive equilibrium cost. H2: prediction of recent game theoretical models; honesty is maintained by condition-dependent trade-offs. Grafen: intersection of H1 and H2. B) Specific predictions derived for each trade-off treatment based on H2.

We derive two conflicting predictions based on previous literature. The Handicap Principle (Zahavi, 1975; Zahavi and Zahavi, 1997) predicts that signals will be honest only if high-quality signallers have to pay a wasteful (positive) cost for sending the high-intensity signal. In contrast, more recent game theoretical models (Hurd, 1995; Számadó, 1999; Lachmann et al., 2001; Bergstrom et al., 2002) predict that signals will be honest under the condition of differential (condition-dependent) trade-offs irrespective of the cost paid at the equilibrium by honest signallers.

H1: The signal cost paid by high-quality signallers maintains the honesty of signals.

H2: The honesty of communication is maintained by signalling trade-offs, in such a way that for honest communication it has to be fulfilled that the marginal cost of high-intensity signal compared to the low-intensity signal for high-quality signallers must be cheaper than for low-quality signallers.

H2 means that the difference in costs between the high-intensity and the low-intensity signal has to be smaller for high-quality signallers than for low-quality signallers. H2 does not exclude the possibility that costs are negative, hence H1 and H2 are in conflict (as seen in Figure 3.A). The implications of Grafen (1990) could be seen as the intersection of H1 and H2 (bottom right cell in Figure 3.A).

We need to elaborate more detailed predictions implied by H2 (see Figure 3.A and B for the experimental design and comparison of predictions). In the condition where the costs of the two signals are equal for both types of signallers – or in other words where there is no trade-off between the signals – high-quality signallers cannot gain anything by selecting one particular signal with the same cost, because costs are the same also for low-quality signallers. Therefore, honest communication cannot evolve since there are no such individual strategies from which it is not beneficial for either party to move towards another action.

H2a: In the absence of trade-off, (a) the signallers are expected to select randomly from the two available signals, therefore (b) receivers respond also randomly to the signals they receive; (c) the accuracy of decisions made by receivers are not better than picking randomly without signals.

In the condition in which there is a fixed trade-off between the two signals, both low- and high-quality signallers are expected to send the same signal. Hence, we expect a pooling equilibrium, in which both signallers send the low-intensity signal and therefore, the receiver is unable to predict the type of the signaller.

H2b: If both types of signallers have the same trade-off between the two signals, then (a) both signaller type will use the same signal, and (b) receivers decide randomly on the allocation of the source; (c) the accuracy of decisions made by receivers are not better off than picking randomly without signals.

Payoffs under the condition-dependent trade-off treatment have been chosen to fit the conditions of separating equilibrium (signaller types separate by sending different signals, i.e. signalling is honest), derived by discrete game theoretical models (Számadó, 1999). That is, the potential benefit of signalling out-weights the potential cost for high-condition signallers, while it is not the case for low-quality signallers.

High-quality signallers will switch to the high-intensity signal in order to acquire the resources from the receiver since in this case the marginal cost of moving to this signal is smaller than the marginal benefit of it. Low-quality signallers will stick to choosing for the low-intensity signal, and therefore honest communication develops.

H2c: If the two signaller types have a different trade-off between the use of two signals in such a way that high-quality signallers have larger marginal benefits for choosing the high-intensity signal than low-quality signallers, then (a) the two types will use different signals depending on their type, thus (b) the receivers will be able to determine the type of signallers correctly and respond differently; (c) the accuracy of decisions made by receivers are better off than picking randomly without signals, i.e. they can allocate the resources according to their interest.

Predictions are summarized in Figure 3.B.

4.3. Methods

Participants played a simple 2x2 signalling game in a computer lab. The experiment took place at the Corvinus University of Budapest (CUB) in Hungary between 25th January 2018 and 15th January 2019. Participants were regular or corresponding students. In total, 12 sessions were organized, involving different numbers of participants (groups of 12, 16, and 20) and 196 students participated in the experiment (118 women out of 196 students). A mix of within and between subject design was applied: within a session, each group played three selected treatments of the nine possible conditions. The selected games varied from session to session and each condition was played as the first, second and third game. Since the order of games and the number of participants (Bruner et al. 2014) may affect the speed of learning dynamics in the game, we control for these factors with statistical methods during the analysis. During the experiment, participants were seated randomly in front of the computers, thus participants took part in the experiment anonymously. Computers were connected on a local network with the help of the software z-Tree (Fischbacher 2007). The description of the game was displayed on participants' screens (see Supplementary Materials Section 5). Instructions were provided on paper as well. At the end of the experiment participants received a show-up fee of HUF 1000 and the payoff of one randomly selected round.

treatment no.	trade-off	Low-quality signaller		High-quality signaller		Expectation by the Handicap Principle	Expectation by game theoretical models
		low-intensity signal	high-intensity signal	low-intensity signal	high-intensity signal		
		c_{LL}	c_{LH}	c_{HL}	c_{HH}		
1.	no	0	0	-100	-100	dishonest	dishonest
2.	yes	0	1500	-1600	-100	dishonest	dishonest
3.	condition-dep.	0	1500	-300	-100	dishonest	honest
4.	no	0	0	0	0	dishonest	dishonest
5.	yes	0	1500	-1500	0	dishonest	dishonest
6.	condition-dep.	0	1500	-200	0	dishonest	honest
7.	no	0	0	100	100	honest	dishonest
8.	yes	0	1500	-1400	100	honest	dishonest
9.	condition-dep.	0	1500	-100	100	honest	honest

Table 1. Treatment conditions. The table shows the costs low- and high-quality signallers have to pay (in HUF) for using low-intensity or high-intensity signals (cLL for low-intensity signal, cLH for high-intensity signal for a low-quality signaller, cHL for low-intensity signal, cHH for high-intensity signal for a high-quality signaller). Costs are indicated with positive numbers. Negative numbers indicate negative costs (i.e., positive rewards for choosing the given signal). First, treatments differed in whether the use of the high-intensity signal is costly, cost-free, or profitable for high-quality signallers. Second, treatments also differed in whether there was no trade-off between sending the low-intensity and high-intensity signal (cLH-cLL=cHH-cHL=0), there was a fixed trade-off between the two signals (cLH-cLL=cHH-cHL=1500), or the trade-off between the two signals was condition-dependent (cLH-cLL=1500; cHH-cHL=200).

4.3.1. Experimental task

In each round, as a first step, participants were randomly divided into groups of four containing two signallers (S) and two receivers (R). In the experiment we called them Player X and Y, respectively. Moreover, we used an unbiased signalling game, where the two types of signallers (high- and low-quality S) were assigned randomly by the computer. In a group of 20, for instance, this meant 10 receivers, 5 high-quality, and 5 low-quality signallers. Groups and roles were changed round-by-round. In the experiment instead of using ‘high’ and ‘low quality’, neutral categories were used to differentiate between the two types. High-quality signallers were presented as blue players and low-quality signallers were depicted in yellow. Everyone was only aware of their own type; they did not know each other’s type.

First, Player X (either blue or yellow) had to choose a signal and send it to Player Y in order to get the resource. High- and low-intensity signals were also replaced by neutral pairs of signs (‘)(’, ‘~’; ‘//’, ‘O’, ‘[]’, ‘<>’). These characters and their costs (that could have been a benefit in the negative costs treatment) in the given condition were displayed. While Player X in blue or yellow condition has selected one of the two displayed characters, Player Y was waiting.

In the next step, receivers (Player Y) decided whether they would give the resource after a signal (the character sent by Player X) was seen. Signallers succeeded if they received the resource, but receivers’ success depended on the (hidden) type of the signaller (high-quality S was preferred to low-quality S). After these steps, participants have learned the type of the sender, the signal they sent and the success of their decision. The game was played for 20 rounds (Blume et al., 1998). In case of a successful decision, they received HUF 1200. In addition to this, the cost of the signal influenced the final payment (see Table 1).

In the interests of clarity, we show an example of the calculation:

Example 1. Y can select from the following signs: (1) using the <> sign reduces Y’s payoff by HUF 300; (2) using the [] sign increases Y’s payoff by HUF 100. Y decided to use the <> sign and X gave him the resource. In this case, the payoff of Y will be 1200 - 300 = 900.

4.4. Analytical Approach

As main outcome measures of our analysis, we defined Honesty, Trust, and Accuracy, and we examine these outcomes in each treatment. First, we report simple averages and correlation statistics and then we provide in-depth regression analysis to test and control for more design elements of the experiment (for instance treatment and time). The latter allows us to deal with the imbalances between experimental sessions as well (slight differences in group size; order of play; the use of different characters as signals). In the following we describe how we specify each outcome in each analysis.

After showing descriptive statistics of the general tendencies regarding the three outcomes visually in Figures 4-6, we start testing each theoretical prediction by calculating Pearson correlation coefficients for each round in each treatment. The Honesty score shows the correlation between signaller type and signal sent; it can vary between -1 (high-quality signallers send low-intensity signals and low-quality types send high-intensity signals) and 1 (high-quality signallers send high-intensity signals and low-quality types send low-intensity signals), where 0 means no correlation (i.e., signalling is random and hence the signal is uninformative). The Trust score shows the correlation between the receiver's decision and the signal received. It is calculated as the Honesty score, where 0 means that the receivers are not relying on the signal in their decisions (i.e., they are giving out the resource randomly). The Accuracy of the decision by the receiver is measured as the correlation between the allocation of the resource and the quality of the signaller: it varies between -1 (meaning that the receiver allocates the resource to low-quality signallers against their interest) and 1 (meaning that the receiver correctly allocates the resource to high quality signallers), where 0 means that receiver allocates the resource randomly. It is important to note that in the "no trade-off" manipulation both signals can develop to be honest because high-intensity signals are not determined by the costs of the signals (both signals are free or have the same costs). Therefore, it is not straightforward how we interpret Honesty or Trust in this treatment. By looking at overall patters over time in Figure 4 we simply report differences in the use of the two signals by the two signaller types, in receivers' action by the two signals. This will only show us if both signals are used by the participants to develop meaning but tells us less about Honesty and Trust. Honesty and Trust in the no trade-off treatment will be the subject of the regression analysis.

Multilevel mixed-effects logistic regressions are used for analysing binary outcomes. Players made binary choices about sending one of the two signals, and when they decided whether to give the resource at the sight of a signal. Success is also a binary outcome, if giving resource to high-quality signallers and refusing it from low-quality players is considered as 1 and any other decision as 0. Specifically, by testing Honesty we estimate whether the probability of selecting low-intensity signals (dependent variable) by high-quality signallers (independent variable) is higher than doing so by low-quality signallers. Negative effect indicates honesty, meaning that compared to low quality signallers the probability of sending low-intensity signals is lower if decision maker is a high-quality signaller. Trust is also translated into probabilities: the probability of refusing the resource (dependent variable) after observing the low-intensity signal (independent variable) compared to observing high-intensity signals. Positive effects indicate more Trust. Accuracy is specified again as the probability of refusing giving the resource but now the condition is the hidden type of the signaller. These models test whether the probability of refusing the resource is higher when a low-quality signaller asks for it. Again, positive parameter estimations imply Accuracy.

In these models we test the temporal dimension of these outcomes by including an interaction effect between the above-mentioned independent variables (Honesty, Trust, Accuracy) and round of the games (see models with a label 'b' in SM Table 17-21). The same sign of the interaction term and the parameter estimation of the main independent variable (the type of the signaller and the signal) indicates that Honesty, Trust and Accuracy are intensifying over time. In addition to this in the models we control for the equilibrium cost manipulations to show that results are independent of it. Lastly, we use control variables to fix imbalances between the experimental sessions. In the text we report the estimated coefficients, and odds ratios as well to indicate the size of the effects.

4.5. Results

In this section three outputs will be examined: (a) to what extent are signals with different intensity separated by the signallers' condition (i.e., whether signals are honest; Honesty score), (b) how resource allocation is made according to which signal was seen (i.e., whether the receiver responds selectively to the signal; Trust score), and (c) the overall degree of coordination between response and condition, to what extent has the quality of the signallers been successfully determined by the receivers (i.e., whether the receivers were successful in making an optimal decision; Accuracy). In the following we provide more precise definitions of these outcomes in accordance with the methodology we use.

	Honesty (%)		Trust (%)		Accuracy (%)	
	dishonest	honest	inconsequential	meaningful	failure	success
negative eq. cost	40.63	59.38	44.44	55.56	45.21	54.79
zero eq. cost	43.04	56.96	44.67	55.33	47.28	52.72
positive eq. cost	39.66	60.34	43.37	56.63	48.46	51.54
Pearson's chi-squared test	4.78		0.77		3.62	
no trade-off	50.92	49.08	49.61	50.39	51.84	48.16
trade-off	44.09	55.91	48.94	51.06	48.22	51.78
conditional trade-off	29.03	70.97	33.64	66.36	41.93	58.07
Pearson's chi-squared test	174.09***		116.64***		33.61***	

** p<0.05, ** p<0.01, *** p<0.001*

Table 2. Honesty, Trust and Accuracy by treatment conditions with statistical tests for significant differences.

4.5.1. Aggregate results

We look at the aggregate results first (see Figures 4 and 6), and then we run more complex models to test the outlined outputs after we control for other variables (signs used in the game, rounds, order of the games number of participants; see tables and detailed descriptions of the effects in Supplementary Materials Section 3 and Section 4).

Figures 4-6 shows the timeline of the experiment as a function of the experimental manipulations (no trade-off, trade-off, conditional trade-off) where three Figures (4-6 respectively) each show: signal use, receiver's decision and receiver's accuracy. Accordingly, honesty, trust and accuracy can be inferred implicitly from the separation of the two lines that are intended to represent the two types of signallers (in Figure 4,6) and signals (Figure 5). Lines approaching the value of 1 shows the increasing proportion of high intensity signals (Figure 4) and resource allocation (Figure 5). Please note that in the “no trade-off” manipulation both signals can function as a high-intensity signal, therefore any preference for one of the two signals in this treatment cannot be interpreted as honesty or trust. Before we address this question in a more detailed assessment, first take a brief look at the simple signal selection in this and in the other trade-off treatments, where costs imply the type of the signal and therefore also honesty and trust.

There is no association between signal use and signaller’s type because both signals are cost-free or have the same positive or negative cost in the “no trade-off” manipulation, meaning that it is not determined which signal would attract receivers’ positive decision. Accordingly, signallers try to assign meaning to both signals, therefore we do not see any difference between the two lines (Figure 4, first column –blue-red). There is some separation of signaller types under the “trade-off” manipulation, because higher proportion of high-quality signallers use high-intensity signals (Figure 4, second column –blue-red). Finally, there is a clear separation of signaller types under the “conditional trade-off” manipulation (Figure 4, third column –blue-red). The receiver’s decisions seem to follow this trend (Figure 5, second columns- green-orange); yet an increase in accuracy is only observed in the last manipulation (Figure 6 third columns, blue-red).

Plotting Honesty and Trust together also separates the manipulations (Figure 7). In Figure 7 Honesty and Trust are expressed as Pearson correlation coefficients between signallers’ state and signals, and signals and actions of the receivers, respectively. Figure 7.A shows the Honesty and Trust scores by manipulations per rounds, while Figure 7.B displays the average of Honesty and Trust scores per manipulations. Average scores indicate that the trust placed into signals is proportional to the honesty of signals. Honesty and Trust scores in the “conditional trade-off” manipulation - irrespectively of the signal cost - are the highest, indicating that signals in this manipulation were the most honest and most trusted. Even in this condition, the values are far from perfect correlations (i.e. far from the top right corner).

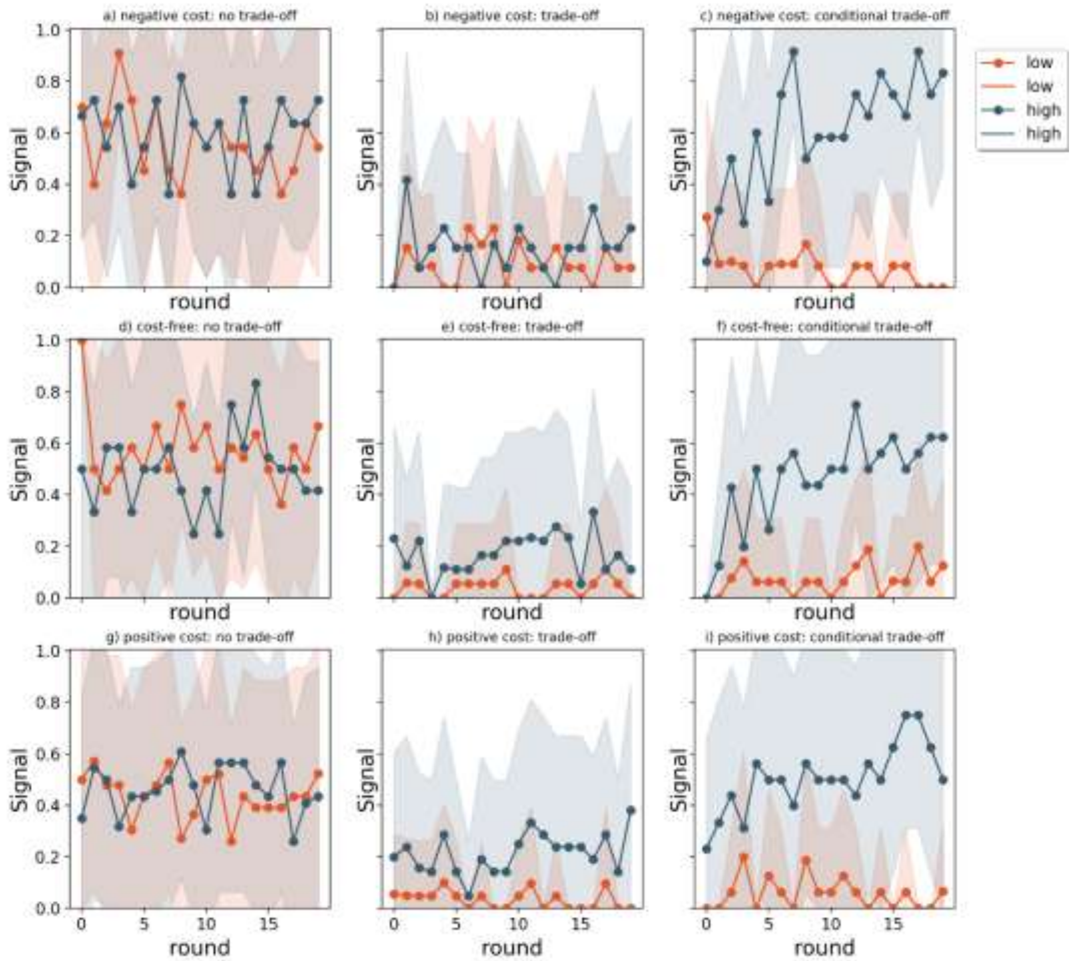


Figure 4. Signaller decisions by rounds and by the nine treatments. Left column: no trade-off manipulation; middle column: trade-off manipulation; right column: conditional trade-off manipulation. Upper row: negative signal cost for high-quality signallers, middle row: cost-free for high-quality signallers, bottom row: positive signal cost for high-quality signallers. 'Signaller decision' panels: dark blue line shows the proportion of high-quality signallers using high-intensity signals, the orangered shows the proportion of low-quality signallers using high-intensity signal. At the honest equilibrium all high-quality signallers should use the high-intensity signal and none of the low-quality ones (dark blue at 1; orangered at zero). Dots show the average per round, overlay shows one standard deviation.

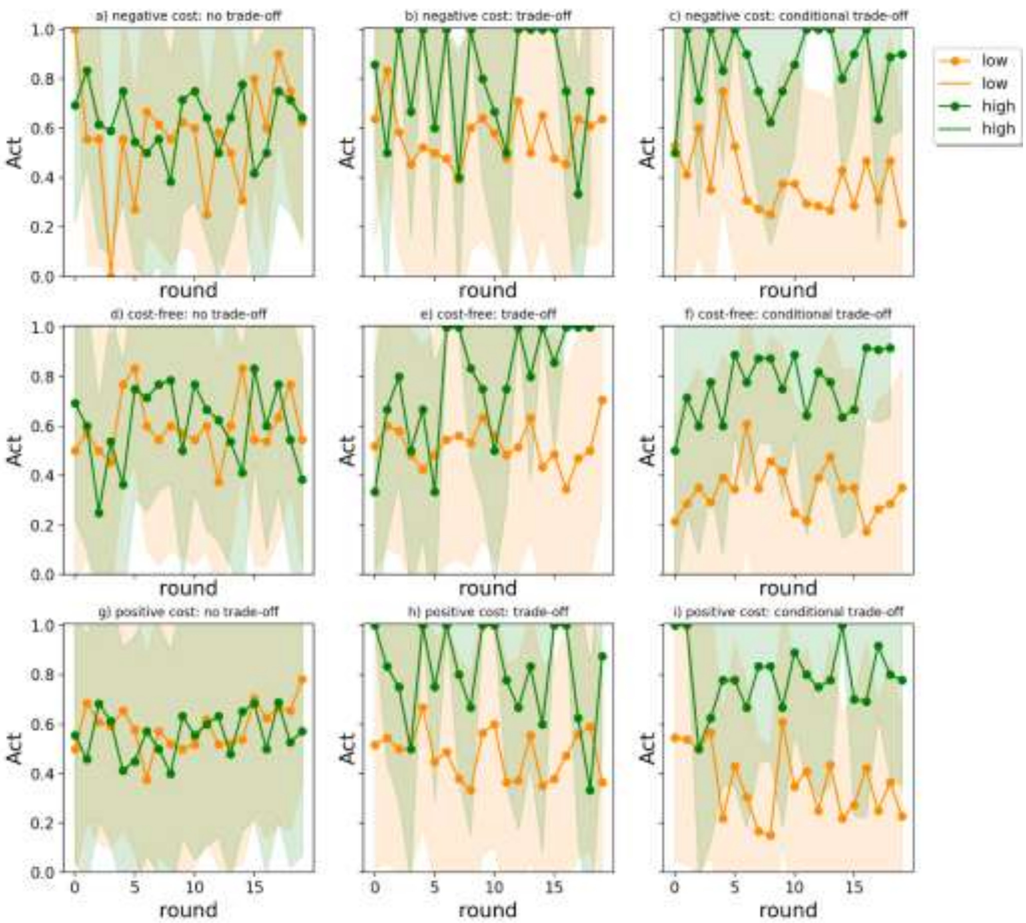


Figure 5. Receiver decisions by rounds and by the nine treatments. Left column: no trade-off manipulation; middle column: trade-off manipulation, right column: conditional trade-off manipulation. Upper row: negative signal cost for high-quality signallers, middle row: cost-free for high-quality signallers, bottom row: positive signal cost for high-quality signallers. 'Receiver decision' panels: green line shows the proportion of receivers giving the resource to high-intensity signals, the orange is for giving the resource to low-intensity signals. At the honest equilibrium receivers should always give the resource in response to high-intensity signals and they should never give to low-intensity ones (green line at 1; orange at 0). Dots show the average per round, overlay shows one standard deviation.

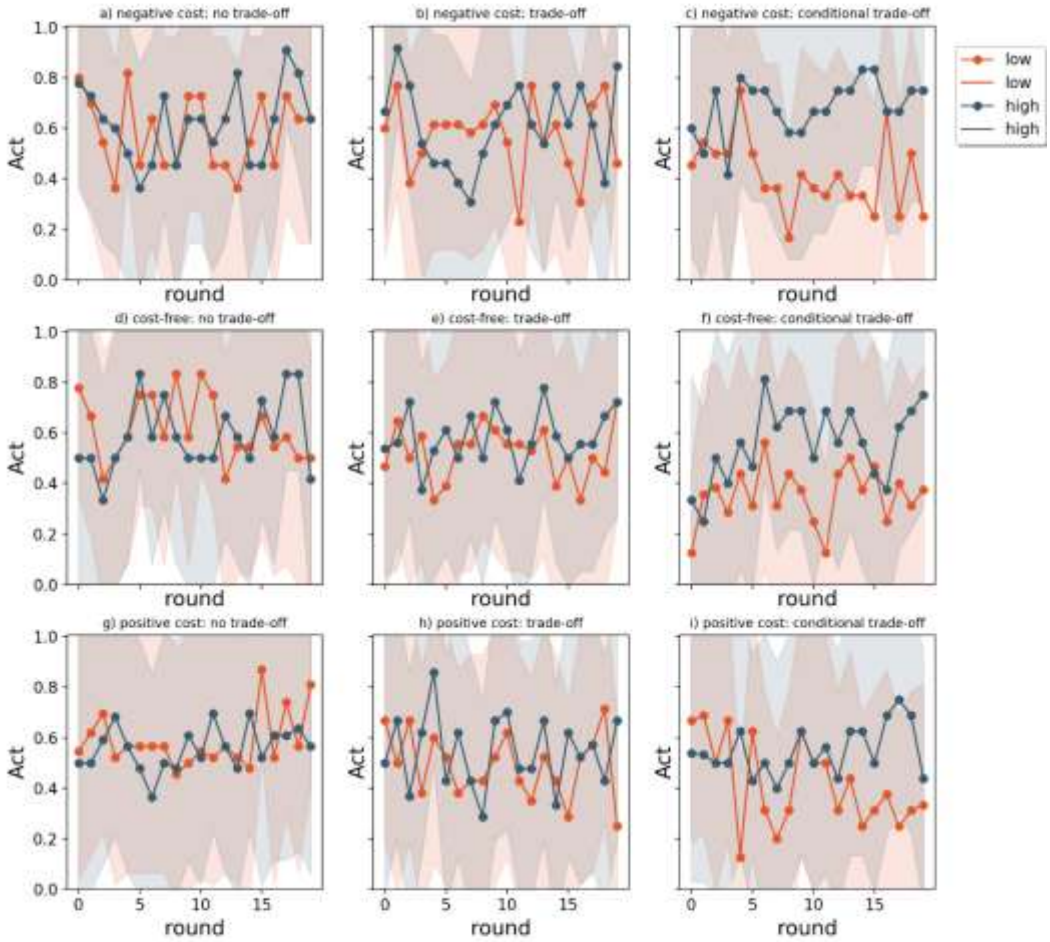


Figure 6. Receivers' accuracy by rounds and by the nine treatments. Left-hand column) No trade-off manipulations, Middle-colum) trade-off manipulations, Right-hand column) conditional trade-off manipulations. Upper row: negative signal cost for high-quality signallers, middle row: cost-free for high-quality signallers, bottom row: positive signal cost for high-quality signallers. 'Receiver accuracy' panels: dark blue and orangered lines show the proportion of High- and Low-quality signallers receiving the resource respectively. At the honest equilibrium, receivers can find out the quality of the signaller from the signal, thus they should donate the resource only to High-quality signallers and never to Low-quality ones (dark blue at 1; orangered at zero). Dots show the average per round, overlay shows one standard deviation.

4.5.2. Hypothesis testing

Next, we go through the results in the order of the hypotheses. Pearson's chi-squared tests are shown to assess our main hypotheses H1 and H2.

H1. We see similar patterns by the manipulation of equilibrium signalling cost (compare results between rows in Figure 4), which suggests that for the signalling cost paid has no effect on honest communication. We cannot see a significant difference between negative, cost-free, and positive cost conditions either in Honesty (χ^2=4.78, p=0.09 see Table 2) or in Trust (χ^2=0.77, p=0.68 see Table 2) and consequently, the same is true for Accuracy (χ^2=3.62, p=0.16 see Table 2).

H2. The honesty of signals, however, is influenced by signalling trade-offs. Figure 4 shows that there are observable differences in both signallers' and receivers' decisions based on trade-off manipulations. Statistical tests confirm these insights: sending signals corresponding to players' status (Honesty) shows variation along trade-off treatments (χ^2=174.09, p<0.001 see Table 2). The meaning of the two signals (Trust) (χ^2=116.64, p<0.001 see Table 2) and the overall coordination between response and condition (Accuracy) diverge as well (χ^2=33.61, p<0.001 see Table 2). In the following, we discuss each treatment in the trade-off manipulation separately by presenting results of regression analyses, where we could handle imbalances between experimental sessions such as number of players, the order of the games, signs used in the experiment.

Mixed-effects logistic regressions are used to test our hypotheses H2a, H2b, and H2c (see detailed results in Supplementary Materials Sections 3 and 4).

H2a. "no trade-off": Different types of signallers used the two signals with the same probability in treatments where there was no difference in the cost of using the two signals (no trade-off) (see the nonsignificant effect of signaller's state on sending separating signals in Model 1a, SM Table 17). Moreover, receivers did not trust one signal more than the other (see non-significant effect of the Signal on Receiver's decision in Model 1a, SM Table 18). At the same time, Accuracy has improved over time, because high-quality signallers were 3.05 percent more likely to receive the resource in each subsequent round (exp(-0.031)), while low-quality signallers were 1.05 times (exp(0.048)) more likely to be rejected by the receiver than high-quality signallers in each round (see Round and Signaller's state – round interaction effects in Model 1b, SM Table 19). The overall Accuracy was 48.16% in this treatment (Table 2), which is not different from random (t(3038)=1.402, p= 0.08). To resolve the apparent contradiction behind Accuracy in the absence of Honesty and Trust, we carried out an analysis on the local conventions that may have developed within each experimental group in the no trade-off manipulation. We found that in half of the groups high-quality signallers used one signal in a higher proportion than low-quality signallers, while the other signal was used by them with a higher probability in the other half of the experimental groups (SM Figure1). We re-coded these signals as high-intensity signals (that were more frequently used by high-quality signallers than low-quality signallers) and labelled the other signal as low-intensity signal. By doing so, we artificially synchronize signals and signallers in this treatment in order to investigate the efficiency of local conventions. We found that local conventions emerged, – again, where the directions of these conventions was random (i.e. they randomly picked one or the other signal as "high-intensity" see SM Figure 1) – because, low-intensity signals are 45.12 percent less likely (exp(-0.600)) to be selected by high-quality signallers than by low-quality signallers (see the effect of Signaller's state in Model 1a, SM Table 20). Initial convention, however, did not intensify over time (see non-significant effect of Signaller's state - Round interaction in Model 1b, SM Table 20).

Looking at the receiver's responses to these signals, we also see a difference in the extent to which the receiver gave the resource after seeing the low-intensity signal, . low-intensity signal were 1.46 times more likely (exp(0.382)) to be rejected by the receiver than high-intensity signals (see the effect of signal intensity in Model 1a, SM Table 21). More interestingly, this separation has developed over time: in each round the probability of refusing the resource upon high-intensity signals decreased by 3.15 percentage (exp(-0.032)) while, in each round, the rejection of low-intensity signals were 1.05 times more likely than the rejection of high-intensity signals (see Round and Signal-Round interaction effect in Model 1b, SM Table 21).

Despite the development of local conventions, in the "no trade-off" treatments receivers estimated signallers' original states correctly only half of the time (48.2%) (gave the resources to high-quality S and refuse to give it to low-quality S), suggesting the observed intention for separation, this treatment fails to provide accuracy that is better than random.

All in all, signals were not informative and receivers did not react selectively on a global level in the no trade-off manipulation, however, local conventions were formed. While the direction of these local conventions was random (in parallel to the aggregate lack of honesty and trust) these conventions were learned by receivers, and as a result the performance of the receiver improved over time. Still, receivers could not do better than random choice when allocating the resource.

H2b. “trade-off”: In treatments with the same trade-off for both type of signallers, we see that the majority (55.91%, Table 2) of signallers were using one signal. Contrary to our expectation, high-quality senders intended to separate signals as they slightly shift from low to high-intensity signals (see the effect of Signaller’s state in Model 2a, SM Table 17). More precisely, selecting low-intensity signals was 83.25 percentage less likely among high-quality signallers than among low-quality signallers (exp(-1.787)). The difference, however, does not increase over time (see Signaler’s state and Round interaction in Model 2b, SM Table 17). We also see a difference within receivers’ decisions by signals: receivers tend to favour high-intensity signals (giving 79.2 percent of the resource) compared to low-intensity signals (51.6%, see SM Table 12). Indeed, the regression analysis suggest that receiver’s preference for high intensity signals is almost 6 times higher (exp(1.780)) (see the effect of Signals in Model 2a, SM Table 18). R’s success ratio was slightly higher than in the previous treatment (51.78%, Table 2, which is not different than random $t(4158)= -1.306$, $p= 0.10$) and we found evidence that low-quality signallers were 1.04 times more likely to be rejected by rounds compared to high-quality signallers (Model 2b, SM Table 19). In sum, in the trade-off condition, different signals were used by different signallers, receivers also acted selectively, and signals were not informative.

H2c. “conditional trade-off”: Results observed in the conditional trade-off manipulation are consistent with our expectations. Signals diverge with the greatest extent by the type of signaller in these treatments. A learning dynamic can be observed in Figure 3. High-quality signallers gradually switched to the use of a signal with favourable trade-off to them – which need not be costly (see insignificant signaller's state – equilibrium cost interaction effect in Model 3c, SM Table 22) – to achieve coordination. Most (~65%) of high-quality signallers used high-intensity signal in the last rounds of the game while more than 90 percent of low-quality signallers were stuck to the low-intensity signal. According to the regression analysis, players, indeed, move towards separating equilibrium: sending low-intensity signals were 96.61 percent less likely among high-quality signallers compared to low-quality signallers (see the effect of signaller's state in Model 3a, SM Table 17). As a response, 80.4 percent of the resource was given if receivers saw the high-intensity signal and almost two-thirds of the low-intensity signals were rejected (63.7%, see SM Table 13). According to the regression analysis, low-intensity signals were ten times more likely to be rejected than high-intensity signals (exp(2.299), see the signal's effect on R's decision in Model 3a, SM Table 18). In terms of overall accuracy (58.07%, which is different than random t(3518)=-6.449, p<0.001), refusing the resource from low-quality signallers were 2 times more likely (exp(0.888)) than refusing it from high-quality signallers (see figures in the last column in Figure 3 and the effect of signaller's (hidden) state on the resource allocation in Model 3a, SM Table 19). The intensified separation of signals (sending low-intensity signals by high-quality signallers were 14.1 percentage less likely round-by-round than the same action by low-quality signallers exp(-0.152), see Model 3b, SM Table 17) and actions over time (in each round refusing a low-intensity signals were 1.07 times more likely than refusing a high-intensity signals , Model 3b, SM Table 18) and the increasing accuracy of the resource allocation (, in each round the probability of refusing the resource from high-quality signallers decreased by 3.63 percentage (exp(-0.037)) while, in each round, the rejection of low-quality signallers were 1.07 times more likely than the rejection of high-quality signallers, exp(0.070) Model 3b, SM Table 19) are demonstrative results of the development of honest communication. All in all, in the conditional trade-off condition, by the last rounds signals were informative, receivers mostly responded selectively, and they achieved a higher success than random choice.

4.6. Conclusion

How can animals and humans communicate is a fundamental question in biology and the social sciences. When senders and receivers have mutual interests, honest communication can develop (e.g., Maynard Smith and Harper, 2003; Bradbury and Vehrencamp, 2011; Barker et al., 2018). Honest communication, however, is much more difficult to be established in situations with conflict of interest, because low-quality senders could mimic high-quality senders by sending the same signal and therefore making the communication uninformative for the receiver. This study intended to make justice between conflicting theoretical accounts that claimed to explain the emergence of honest communication in situations with conflict of interest by conducting experiments with human participants. The conflicting predictions tested were made by the proponents of the Handicap Principle (Zahavi 1975; Grafen, 1990; Zahavi and Zahavi, 1999) and by subsequent game theoretic models (Hurd, 1995; Getty, 1998; Számadó, 1999; Lachmann et al., 2001; Bergstrom et al., 2002). These implications have been translated into two main hypotheses. First, we tested following the Handicap Principle (Zahavi 1975; Zahavi and Zahavi, 1999) if signalling costs for high-quality signallers alone could establish honest communication (H1). Second, we tested whether signalling trade-offs could lead to honest signalling (H2; see Számadó, 1999; 2011; Lachmann et al., 2001; Bergstrom et al., 2002; Getty, 2006). We expected that the lack of trade-offs (H2a) and trade-offs that are identical for low- and high-quality signallers (H2b) would not be sufficient for receivers to differentiate signaller types efficiently. Instead, condition-dependent trade-offs have been expected to be the guarantees of honest communication (H2c).

Our results support the last claim: informative signals emerged in our experiment under the condition-dependent trade-off condition regardless of the cost of signals for high-quality signallers. Signals with zero or even negative production cost (benefit) were honest if trade-offs were condition-dependent. An intention to separate signaller types by using different signals can be observed in the other manipulations as well (no trade-off, trade-off). Local conventions emerged in the no trade-off condition, although the direction of these conventions were random, thus at the aggregate level no honesty or trust can be observed. There was an intention to separate in the trade-off condition as well, however, it was only mildly successful. Thus, we can see that the intention to separate signaller types (be honest) was successful only in the condition-dependent trade-off manipulations as predicted by game-theoretical models. In other words, full-fledged honesty could not evolve in the no trade-off and trade-off treatments despite the effort of the participants to create an honest system. Potential motivations for this intention (to be honest) will be discussed later.

Our results demonstrate that the conditions that allow the emergence of honest signalling is different and much wider from what was predicted by the Handicap Principle (Zahavi 1975). Equilibrium signalling cost in itself is not enough to generate an honest signalling system. Also, the region of honesty is not constrained to the manipulation with differential marginal cost with positive equilibrium cost as predicted by Grafen (1990). Last but not least, the highest level of honesty was achieved under the condition-dependent trade-off condition with negative signal cost contrary to the prediction of the Handicap Principle.

Not surprisingly, there is an increasing gap between the predictions of the Handicap Principle and the empirical observations. There is a growing body of literature showing the abundance of dishonest signals both in nature (Backwell et al., 2000; Christy and Rittschof, 2011; Brown, Garwood and Williamson, 2012; Dalziell et al., 2016; Casewell et al., 2017; Wang et al., 2017; Fujisawa et al., 2020) and in human communication (DePaulo et al., 1996; Kashy and DePaulo, 1996; Hess and Hagen, 2006; Fischbacher and Föllmi-Heusi, 2013, Abeler et al. 2019), in parallel with a growing literature showing that signals are not costly to produce (Borgia, 1993, 1996; McCarty, 1996; Moreno-Rueda, 2007; McCullough and Emlen, 2013; Askew, 2014; Thavarajah et al., 2016; Guimarães et al., 2017). For example, recent studies found that the flagship example of the Handicap Principle, the peacock's train, does not handicap the locomotion of peacocks (Askew, 2014; Thavarajah et al., 2016). Also, there is a growing recognition that cheap (subtle) signals can be lot more important in human communication as it would be expected based on the Handicap Principle (Bliege Bird and Power, 2015; Bliege Bird et al., 2018). Last but not least, these signalling trade-offs can be implemented between benefit functions, i.e., having a cost function per se is not a requirement of honest signalling. Our results thus reinforce the call for the replacement of the Handicap Principle with a Darwinian theory of signalling based on conditional signalling trade-offs (Getty, 2006; Penn and Számadó, 2020) for explaining honest communication.

Overall, it is important to emphasize that no strict separating equilibria were observed in any of the treatment conditions. There was always some level of mixing indicating some randomness of decisions even among our subjects who were university students well able to grasp the structure of the simple experimental signalling game. The level of mixing, however, was much reduced under the condition-dependent trade-off manipulation. Some differentiation evolved between low- and high-type signallers in the simple trade-off manipulation. Only the condition-dependent trade-off manipulation allowed receivers to make an informed decision (i.e. do better than random choice). There could be several factors influencing the level of mixing observed in the experiment. The first one is the length of the experimental sessions. As a result of a learning process, a higher differentiation is expected between signaller types in the use of signal over time. The equilibrium signaller and receiver strategies were straightforward, yet they were somewhat contra-intuitive with negative signal cost. In these later manipulations, it was naturally a longer learning process to find (or being close to) the equilibrium strategy pair. The second factor is the ratio of marginal benefit to marginal cost. The higher this ratio, the more obvious is the marginal benefit, thus probably it is easier to find (or get closer to) the honest equilibrium (if there is any) in such situations. The third factor is the roles played by the participants. Participants alternated between all three roles in our experiment (i.e., receiver, low- and high-quality signaller). There are advantages and disadvantages of this setup. While participants might have a better understanding of the experimental setup, they may not be playing (optimizing for) a single role. We observed an intention to separate (to signal honestly) in all manipulations, which could be due to the fact that players alternated between all three roles, which in turn created an incentive to be honest (since both receivers and high-quality signallers are better off this way). Playing a single role might force them to optimize their actions for that single role only, thus this could remove the above effect. Alternatively, this observed effect could be due to a more general ‘truth telling’ tendency identified in previous ‘honesty’ experiments (i.e. dice roll experiments refs …, see Abelar et al., 2018 for review). The last potential factor, unaddressed in this experiment, is the potential role of individual variation. Both previous experiments on cooperation and the famous ‘dice roll’ experiments revealed an heterogeneity of human predispositions: some of us are more cooperative or more willing to tell the truth than the others and vica versa. If so, achieving perfect separation (i.e., a perfectly honest equilibrium) may not be possible at all. Only future experiments can tell which one these factors are responsible for the observed level of mixing in our results.

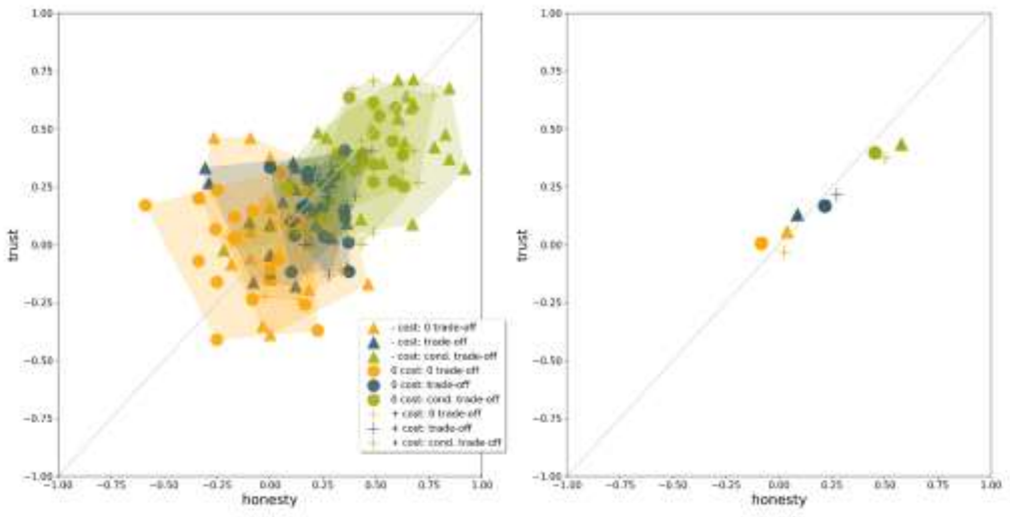


Figure 7. Honesty and Trust scores as a function of experimental manipulations. A) Honesty and Trust scores by rounds and by manipulations, polygons show the rounds belonging to the same manipulation (left); B) the Honesty and Trust score of each manipulation (right). Orange: no trade-off, Blue: trade-off, Green: conditional trade-off; Circle: negative cost, Triangle: no cost, Square: positive cost.

4.7. Supplementary

Section 1. Summary table of the experimental groups

Table 1 Design elements of experimental groups

treatment no.	cost of high-intensity signal	trade-off	Played before honest eq.	Played after honest eq.	Total nr. of participant	% of female participants	Average age
1	Negative	no	12 (‘O’ ‘//’)	16 (‘~’ ‘)(‘),16 (‘<>’ ‘[]’)	44	0.591	24.9
2	Negative	yes	16 (‘O’ ‘//’), 16 (‘~’ ‘)(‘), 20 (‘~’ ‘)(‘)	-	52	0.673	23.5
3	Negative	condition-dep.	16 (‘O’ ‘//’), 16 (‘~’ ‘)(‘), 16 (‘<>’ ‘[]’)		48	0.646	23.8
4	Zero	no	16 (‘<>’ ‘[]’)	16 (‘O’ ‘//’), 16 (‘~’ ‘)(‘)	48	0.604	24.8
5	Zero	yes	16 (‘~’ ‘)(‘), 20 (‘O’ ‘//’)	16 (‘<>’ ‘[]’), 20 (‘~’ ‘)(‘)	72	0.611	23.1
6	Zero	condition-dep.	12 (‘<>’ ‘[]’), 16 (‘O’ ‘//’), 16 (‘~’ ‘)(‘), 20 (‘O’ ‘//’)		64	0.484	22.6
7	Positive	no	12 (‘~’ ‘)(‘), 16 (‘<>’ ‘[]’)	12 (‘O’ ‘//’), 20 (‘~’ ‘)(‘)	60	0.633	22.4
8	Positive	yes	12 (‘<>’ ‘[]’), 16 (‘<>’ ‘[]’)	16 (‘O’ ‘//’), 20 (‘<>’ ‘[]’), 20 (‘O’ ‘//’)	84	0.595	22.2
9	Positive	condition-dep.	12 (‘~’ ‘)(‘), 16 (‘<>’ ‘[]’), 16 (‘O’ ‘//’), 20 (‘<>’ ‘[]’)		64	0.578	23.8

Table 1 shows the number of players played and the sign they used in each experimental group. We tried to randomly achieve variability in group size, order, and signals used by randomizing these characteristics between treatments.

Section 1.

Table 2a Honesty by equilibrium cost treatments

	negative eq. cost (%)	zero eq. cost (%)	positive eq. cost (%)	Total (%)
dishonest	40.63	43.04	39.66	41.08
honest	59.38	56.96	60.34	58.92
Total	100	100	100	100

Pearson chi2(2) = 4.7782 Pr = 0.092

Table 2b Local conventions by equilibrium cost treatments

	negative eq. cost (%)	zero eq. cost (%)	positive eq. cost (%)	Total (%)
dishonest	38.54	40.33	38.61	39.18
honest	61.46	59.67	61.39	60.82
Total	100	100	100	100

Pearson chi2(2) = 1.5483 Pr = 0.461

Table 3a Honesty by trade-off treatments

	no trade-off (%)	trade-off (%)	conditional trade-off (%)	Total (%)
dishonest	50.92	44.09	29.03	41.08
honest	49.08	55.91	70.97	58.92
Total	100	100	100	100

Pearson chi2(2) = 174.0946 Pr = 0.000

Table 3b Local conventions by trade-off treatments

	no trade-off (%)	trade-off (%)	conditional trade-off (%)	Total (%)
dishonest	44.21	44.09	29.03	39.18
honest	55.79	55.91	70.97	60.82
Total	100	100	100	100

Pearson chi2(2) = 113.1865 Pr = 0.000

Table 4a Trust by equilibrium cost treatments

	negative eq. cost (%)	zero eq. cost (%)	positive eq. cost (%)	Total (%)
inconsequential	44.44	44.67	43.37	44.1
meaningful	55.56	55.33	56.63	55.9
Total	100	100	100	100

Pearson chi2(2) = 0.7704 Pr = 0.680

Table 4b Trust by equilibrium cost treatments with local conventions in the no-trade off treatment

	negative eq. cost (%)	zero eq. cost (%)	positive eq. cost (%)	Total (%)
inconsequential	44.17	45.11	42.31	43.77
meaningful	55.83	54.89	57.69	56.23
Total	100	100	100	100

Pearson chi2(2) = 3.239 Pr = 0.198

Table 5a Trust by trade-off treatments

	no trade-off (%)	trade-off (%)	conditional trade-off (%)	Total (%)
inconsequential	49.61	48.94	33.64	44.1
meaningful	50.39	51.06	66.36	55.9
Total	100	100	100	100

Pearson chi2(2) = 116.6368 Pr = 0.000

Table 5b Trust by trade-off treatments with local conventions in the no-trade off treatment

	no trade-off (%)	trade-off (%)	conditional trade-off (%)	Total (%)
inconsequential	48.42	48.94	33.64	43.77
meaningful	51.58	51.06	66.36	56.23
Total	100	100	100	100

Pearson chi2(2) = 109.4042 Pr = 0.000

Table 6 Accuracy by equilibrium cost treatments

	negative eq. cost (%)	zero eq. cost (%)	positive eq. cost (%)	Total (%)
failure	45.21	47.28	48.46	47.18
success	54.79	52.72	51.54	52.82
Total	100	100	100	100

Pearson chi2(2) = 3.6248 Pr = 0.163

Table 7 Accuracy by trade-off treatments

	no trade-off (%)	trade-off (%)	conditional trade-off (%)	Total (%)
failure	51.84	48.22	41.93	47.18
success	48.16	51.78	58.07	52.82
Total	100	100	100	100

Pearson chi2(2) = 33.6142 Pr = 0.000

Table 8a Signals by signallers in the no trade-off treatment

	high-quality signaller	low-quality signaller	Total
high-intensity signal	388 51.87	388 52.01	776 51.94
low-intensity signal	360 48.13	358 47.99	718 48.06
Total	748 100	746 100	1494 100

Pearson chi2(2) = 0.0029   Pr = 0.957

Table 8b Signals by signallers in the no trade-off treatment with local conventions

	high-quality signaller	low-quality signaller	Total
high-intensity signal	432 57.75	330 44.24	762 51.00
low-intensity signal	316 42.25	416 55.76	732 49.00
Total	748 100	746 100	1494 100

Pearson chi2(2) = 27.3121   Pr =0.000

Table 9 Signals by signallers in the trade-off treatment

	high-quality signaller	low-quality signaller	Total
high-intensity signal	192 18.70	56 5.45	248 12.07
low-intensity signal	835 81.30	971 94.55	1806 87.93
Total	1027 100	1027 100	2054 100

Pearson chi2(2) = 84.8221   Pr = 0.000

Table 10 Signals by signallers in the conditional trade-off treatment

	high-quality signaller	low-quality signaller	Total
high-intensity signal	451 52.26	58 6.78	509 29.61
low-intensity signal	412 47.74	798 93.22	1210 70.39
Total	863 100	856 100	1719 100

Pearson chi2(2) = 426.5519   Pr = 0.000

Table 11a Acts by signals in the no trade-off treatment

	high-intensity signal	low-intensity signal	Total
Give	454 58.66	403 56.63	857 57.56
Refuse	320 41.34	312 43.64	632 42.44
Total	774 100	715 100	1489 100

Pearson chi2(2) = 0.7997   Pr = 0.371

Table 11b Acts by signals in the no trade-off treatment with local conventions

	high-intensity signal	low-intensity signal	Total
Give	456 60.00	401 55.01	857 57.56
Refuse	304 40.00	328 44.99	632 42.44
Total	760 100	729 100	1489 100

Pearson chi2(2) = 3.7974   Pr = 0.051

Table 12 Acts by signals in the trade-off treatment

	high-intensity signal	low-intensity signal	Total
Give	190 79.17	929 51.58	1119 54.83
Refuse	50 20.83	872 48.42	922 45.17
Total	240 100	1801 100	2041 100

Pearson chi2(2) = 65.0621 Pr = 0.000

Table 13 Acts by signals in the conditional trade-off treatment

	high-intensity signal	low-intensity signal	Total
Give	406 80.40	435 36.34	841 49.41
Refuse	99 19.60	762 63.66	861 50.59
Total	505 100	1197 100	1702 100

Pearson chi2(2) = 275.7656 Pr = 0.000

Table 14 Acts by signallers in the no trade-off treatment

	high-quality signaller	low-quality signaller	Total
Give	422 56.72	435 58.39	857 57.56
Refuse	322 43.28	310 41.61	632 42.44
Total	744 100	745 100	1489 100

Pearson chi2(2) = 0.4244 Pr = 0.515

Table 15 Acts by signallers in the trade-off treatment

	high-quality signaller	low-quality signaller	Total
Give	588 57.59	531 52.06	1119 54.83
Refuse	433 42.41	489 47.94	922 45.17
Total	1021 100	1020 100	2041 100

Pearson chi2(2) = 6.3043 Pr = 0.012

Table 16 Acts by signallers in the conditional trade-off treatment

	high-quality signaller	low-quality signaller	Total
Give	507 59.44	334 39.34	1119 49.41
Refuse	346 40.56	515 60.66	922 50.59
Total	1021 100	1020 100	2041 100

Pearson chi2(2) = 68.7502   Pr = 0.000

Section 3. Regression tables

Table 17 Results of the multilevel mixed-effects logistic regressions, Honesty

Dependent variable: signals (1: low-intensity signal 0: high-intensity signal)	Model 1a no trade-off	Model 1b no trade-off	Model 1c no trade-off	Model 2a trade-off	Model 2b trade-off	Model 3a cond. trade-off	Model 3b cond. trade-off
Signaller’s state							
baseline: low-quality							
high-quality	-0.019	0.145	0.073	-1.787***	-1.207**	-3.386***	-1.750***
	(-0.17)	-0.62	(0.60)	(-9.57)	(-3.27)	(-17.69)	(-5.05)
Round	0.015	0.022	0.015	0.001	0.040	-0.080***	0.031
	(1.55)	-1.66	(1.44)	(0.09)	(1.53)	(-6.16)	(1.22)
Signaller’s state – round interaction							
high-quality * round		-0.015			-0.056		-0.152***
		(-0.80)			(-1.77)		(-5.06)
Equilibrium cost							
baseline: negative							
cost-free	0.388*	0.388*	0.120	0.335	0.323	0.486	0.505
	(2.33)	-2.33	(0.37)	(1.15)	(1.11)	(1.60)	(1.61)
positive	0.744***	0.745***	0.275	0.489	0.453	0.548	0.575
	(4.12)	-4.12	(0.79)	(1.18)	(1.09)	(1.83)	(1.86)
Played after honest equilibrium	0.416	0.416	0.345	-0.515	-0.498		
	(1.81)	-1.8	(0.73)	(-1.17)	(-1.13)		
Nr. of players	-0.075	-0.075	0.043	-0.098	-0.099	0.045	0.048
	(-1.74)	(-1.74)	(0.73)	(-1.13)	(-1.15)	(0.76)	(0.78)
Signs							
baseline: <>[]							
O//	-0.531*	-0.531*	-0.071	-0.366	-0.363	0.553*	0.573*
	(-2.10)	(-2.09)	(-0.13)	(-1.34)	(-1.33)	(1.99)	(2.00)
~)(	-0.233	-0.233	-0.339	-0.012	-0.029	-0.197	-0.207
	(-1.19)	(-1.19)	(-1.04)	(-0.04)	(-0.09)	(-0.74)	(-0.76)
Constant	0.480	0.4	-1.237	5.645***	5.292***	2.894**	1.632
	(0.73)	-0.6	(-1.50)	(3.73)	(3.46)	(2.85)	(1.53)
random intercept variance between subjects	0.358**	0.360**	0.396**	2.992***	3.004***	1.550***	1.676***
	(3.24)	-3.24	(3.26)	(4.29)	(4.29)	(4.60)	(4.63)
Log likelihood	-1008.04	-1007.723	-852.37	-604.40	-602.83	-722.88	-709.79
Observations	1494	1494	1262	2054	2054	1719	1719

t statistics in parentheses, * *p*<0.05, ** *p*<0.01, *** *p*<0.001

Table 18 Results of the multilevel mixed-effects logistic regressions, Trust

Dependent variable: R's decision (1: refuse 0: give)	Model 1a no trade-off	Model 1b no trade-off	Model 2a trade-off	Model 2b trade-off	Model 3a cond. trade-off	Model 3b cond. trade-off
Signal						
baseline: High-intensity signal						
Low-intensity signal†	0.143	0.332	1.780***	1.217**	2.299***	1.455***
	(1.18)	(1.31)	(9.23)	(3.11)	(15.59)	(4.61)
Round	-0.007	0.001	0.005	-0.042	0.020*	-0.038
	(-0.69)	(0.08)	(0.50)	(-1.39)	(1.99)	(-1.68)
Signal intensity – round interaction						
Low-intensity signal * round		-0.018		0.052		0.075**
		(-0.85)		(1.61)		(2.92)
Equilibrium cost						
baseline: negative						
cost-free	-0.062	-0.063	0.035	0.030	0.141	0.155
	(-0.31)	(-0.31)	(0.21)	(0.18)	(0.62)	(0.67)
positive	0.170	0.174	0.130	0.116	0.321	0.313
	(0.77)	(0.79)	(0.55)	(0.49)	(1.43)	(1.39)
Played after honest equilibrium	-0.330	-0.330	0.503	0.517		
	(-1.02)	(-1.02)	(1.81)	(1.85)		
Nr. of players	-0.007	-0.006	-0.055	-0.058	0.000	-0.004
	(-0.11)	(-0.10)	(-1.04)	(-1.07)	(0.00)	(-0.09)
Signs						
baseline: <>[]						
O//	0.297	0.295	0.110	0.111	0.302	0.293
	(0.91)	(0.91)	(0.62)	(0.62)	(1.45)	(1.40)
~)(	0.113	0.107	0.098	0.093	0.224	0.196
	(0.46)	(0.43)	(0.51)	(0.48)	(1.07)	(0.93)
Constant	-0.220	-0.313	-1.280	-0.725	-2.218**	-1.463
	(-0.24)	(-0.34)	(-1.38)	(-0.73)	(-2.96)	(-1.85)
random intercept variance between subjects	1.147*** (4.59)	1.151*** (4.59)	1.315*** (5.60)	1.329*** (5.60)	0.779*** (4.53)	0.783*** (4.53)
Log likelihood	-934.29	-933.93	-1228.83	-1230.12	-992.52	-988.23
Observations	1489	1489	2041	2041	1702	1702

t statistics in parentheses, * p<0.05, ** p<0.01, *** p<0.001

Table 19 Results of the multilevel mixed-effects logistic regressions, Accuracy

Dependent variable: R's decision (1: refuse 0: give)	Model 1a no trade-off	Model 1b no trade-off	Model 2a trade-off	Model 2b trade-off	Model 3a cond. trade-off	Model 3b cond. trade-off
Signaller's (hidden) state						
baseline: high -quality						
low-quality	-0.145	-0.658**	0.194	-0.230	0.888***	0.143
	(-1.21)	(-2.58)	(1.90)	(-1.07)	(8.24)	(0.63)
Round	-0.007	-0.031*	0.002	-0.018	-0.003	-0.037**
	(-0.67)	(-2.10)	(0.22)	(-1.44)	(-0.27)	(-2.78)
Signaller's state – round interaction						
low-quality * round		0.048*		0.040*		0.070***
		(2.29)		(2.24)		(3.67)
Equilibrium cost						
baseline: negative						
cost-free	-0.044	-0.046	0.053	0.057	0.248	0.251
	(-0.22)	(-0.23)	(0.32)	(0.35)	(1.28)	(1.30)
positive	0.201	0.198	0.142	0.143	0.379*	0.387*
	(0.92)	(0.90)	(0.62)	(0.62)	(1.99)	(2.02)
Played after honest equilibrium	-0.310	-0.313	0.412	0.414		
	(-0.96)	(-0.96)	(1.59)	(1.59)		
Nr. of players	-0.010	-0.011	-0.058	-0.059	0.012	0.011
	(-0.16)	(-0.18)	(-1.17)	(-1.18)	(0.32)	(0.30)
Signs						
baseline: <>[]						
O//	0.273	0.272	0.029	0.036	0.394*	0.400*
	(0.84)	(0.83)	(0.17)	(0.21)	(2.21)	(2.24)
~)(	0.095	0.099	0.057	0.059	0.128	0.129
	(0.39)	(0.40)	(0.31)	(0.32)	(0.71)	(0.71)
Constant	-0.053	0.217	-1.280	0.580	-1.004	-0.629
	(-0.06)	(0.23)	(-1.38)	(0.67)	(-1.59)	(-0.99)
random intercept variance between subjects	1.146***	1.184***	1.315***	1.130***	0.495***	0.493***
	(4.59)	(4.61)	(5.60)	(5.50)	(4.02)	(3.99)
Log likelihood	-934.26	-931.63	-1279.66	-1277.15	-1113.49	-1106.70
Observations	1489	1489	2041	2041	1702	1702

t statistics in parentheses, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 20 Results of the multilevel mixed-effects logistic regressions with local conventions in the no trade-off treatment

Dependent variable†: signals (1: potential low-intensity signal 0: potential high-intensity signal)	Model 1a no trade-off	Model 1b no trade-off	Model 1c no trade-off
Signaller's state			
baseline: low-quality			
high-quality	-0.600***	-0.466*	-0.799***
	(-5.31)	(-1.98)	(-4.20)
Round	0.009	0.015	0.009
	(0.94)	(1.12)	(0.95)
Signaller's state – round interaction			
high-quality * round		-0.013	
		(-0.65)	
Equilibrium cost			
baseline: negative			
cost-free	0.111	0.111	0.116
	(0.65)	(0.66)	(0.68)
positive	-0.062	-0.062	-0.059
	(-0.34)	(-0.34)	(-0.32)
Played after honest equilibrium	0.576*	0.575*	0.429
	(2.46)	(2.45)	(1.65)
Signaller's state - Played after honest equilibrium interaction			0.310
high-quality * Played after honest equilibrium			(1.31)
Nr. of players	0.010	0.011	0.010
	(0.24)	(0.24)	(0.23)
Signs			
baseline: <>[]			
O//	-0.283	-0.282	-0.290
	(-1.10)	(-1.10)	(-1.13)
~)(	-0.030	-0.030	-0.035
	(-0.16)	(-0.15)	(-0.18)
Constant	-0.276	-0.343	-0.178
	(-0.41)	(-0.50)	(-0.26)
random intercept variance between subjects	0.394***	0.394***	0.395***
	(3.48)	(3.48)	(3.48)
Log likelihood	-986.24	-986.02	-985.38
Observations	1494	1494	1494

t statistics in parentheses, * $p<0.05$, ** $p<0.01$, *** $p<0.001$, † See detailed calculation of the dependent variable in Section 4

Table 21 Results of the multilevel mixed-effects logistic regressions with local conventions in the no trade-off treatment

Dependent variable: R's decision (1: refuse 0: give)	Model 1a no trade-off	Model 1b no trade-off
Signal		
baseline: High-intensity signal		
Low-intensity signal†	0.382**	-0.135
	(-3.12)	(0.53)
Round	-0.008	-0.032*
	(-0.77)	(-2.17)
Signal intensity – round interaction		
Low-intensity signal * round		0.049*
		(-2.32)
Equilibrium cost		
baseline: negative		
cost-free	-0.048	-0.041
	(-0.24)	(-0.20)
positive	0.221	0.224
	(1.01)	(1.02)
Played after honest equilibrium	-0.354	-0.334
	(-1.08)	(-1.02)
Nr. of players	-0.010	-0.013
	(-0.17)	(-0.22)
Signs		
baseline: <>[]		
O//	0.291	0.269
	(0.89)	(0.82)
~)(	0.082	0.062
	(0.33)	(0.25)
Constant	-0.275	0.014
	(-0.30)	(0.01)
random intercept variance between subjects	1.177***	1.162***
	(4.61)	(4.59)
Log likelihood	-930.07	-927.37
Observations	1489	1489

t statistics in parentheses, * *p*<0.05, ** *p*<0.01, *** *p*<0.001, † See detailed calculation of the dependent variable in Section 4

Table 22 Results of the multilevel mixed-effects logistic regressions, 3-way interactions

Dependent variable: signals (1: low-intensity signal 0: high-intensity signal)	Model1c no trade-off[†]	Model1d no trade-off[†]	Model 2c trade-off	Model 2d trade-off	Model 3c cond. trade-off	Model 3d cond. trade-off
Signaller's state						
baseline: low-quality						
high-quality	-0.501*	-0.218	-1.972***	-1.811**	-2.940***	-1.934**
	(-2.53)	(-0.52)	(-5.96)	(-2.64)	(-10.27)	(-3.08)
Round	0.009	0.018	0.001	0.016	-0.080***	-0.042
	(0.94)	(0.73)	(0.09)	(0.33)	(-6.15)	(-1.02)
Signaller's state – round interaction						
high-quality * round		-0.027		-0.017		-0.092
		(-0.76)		(-0.29)		(-1.88)
Equilibrium cost						
baseline: cost-free						
negative	0.033	0.163	-1.304**	-1.622*	0.041	-1.574*
	(0.15)	(0.38)	(-3.13)	(-2.10)	(0.09)	(-2.03)
positive	-0.153	-0.181	0.791	0.188	0.588	-0.271
	(-0.68)	(-0.44)	(1.73)	(0.23)	(1.46)	(-0.35)
Signaler's state – equilibrium cost interaction						
high-quality * negative	-0.293	-0.349	1.471**	1.499	-0.715	1.152
	(-1.02)	(-0.58)	(3.24)	(1.60)	(-1.65)	(1.33)
high-quality * positive	-0.035	-0.377	-0.823	0.358	-0.704	-0.450
	(-0.13)	(-0.67)	(-1.72)	(0.38)	(-1.67)	(-0.52)
Round – equilibrium cost interaction						
round * negative		-0.012		0.028		0.170*
		(-0.35)		(0.45)		(2.57)
round * positive		0.002		0.063		0.082
		(0.08)		(0.88)		(1.33)
Round – equilibrium cost - Signaler's state interaction						
round * negative * high-quality		0.005		-0.001		-0.202**
		(0.11)		(-0.02)		(-2.58)
round * positive * high-quality		0.032		-0.118		-0.025
		(0.68)		(-1.42)		(-0.35)
Played after honest equilibrium	0.583*	0.577*	-0.528	-0.531		
	(2.48)	(2.45)	(-1.17)	(-1.16)		
Nr. of players	0.009	0.010	-0.097	-0.100	0.040	0.043
	(0.21)	(0.23)	(-1.09)	(-1.11)	(0.66)	(0.68)
Signs						
baseline: <>[]						
O//	-0.295	-0.289	-0.469	-0.463	0.533	0.559
	(-1.14)	(-1.12)	(-1.65)	(-1.62)	(1.90)	(1.92)
~)(	-0.034	-0.031	-0.085	-0.074	-0.227	-0.224
	(-0.18)	(-0.16)	(-0.27)	(-0.23)	(-0.85)	(-0.81)
Constant	-0.195	-0.298	6.197***	6.136***	3.155**	2.683*
	(-0.26)	(-0.38)	(4.04)	(3.78)	(3.02)	(2.28)
random intercept variance between subjects	0.395***	0.397***	3.174***	3.314***	1.587***	1.744***
	(3.48)	(3.48)	(4.32)	(4.32)	(4.60)	(4.63)
Log likelihood	-985.62	-984.41	-591.56	-587.30	-720.98	-701.47
Observations	1494	1494	2054	2054	1719	1719

t statistics in parentheses, * p<0.05, ** p<0.01, *** p<0.001, [†] See detailed calculation of the dependent variable in Section 4

Section 4. Detailed analysis of regression results

H2a

In treatments without trade-off there is no divergence in signalling between the two types of signallers (see the effect of Signaller's state in Model 1a, Table 17). Since the cost of the two signals was identical, players could converge to two different separating equilibrium. We studied possible directions of separation in each experimental group by looking at the prevalence of the two signals among both type of signallers. We treated any difference in signal use between high-quality and low-quality players as a possible way of separation (see description of Figure 1). By doing so, we artificially create divergence in signalling between the two types of signallers ($\beta=-0.600$, $p<0.001$, Model 1a, Table 20). Players found local conventions independently whether the game was played after a conditional trade-off treatment (see Model 1c in Table 20). The separation towards local conventions, however, did not increase over time (see Signaller's state – round interaction, Model 1b in Table 20). We do not see differences in these conventions between the equilibrium cost treatments (Model 1d in Table 22).

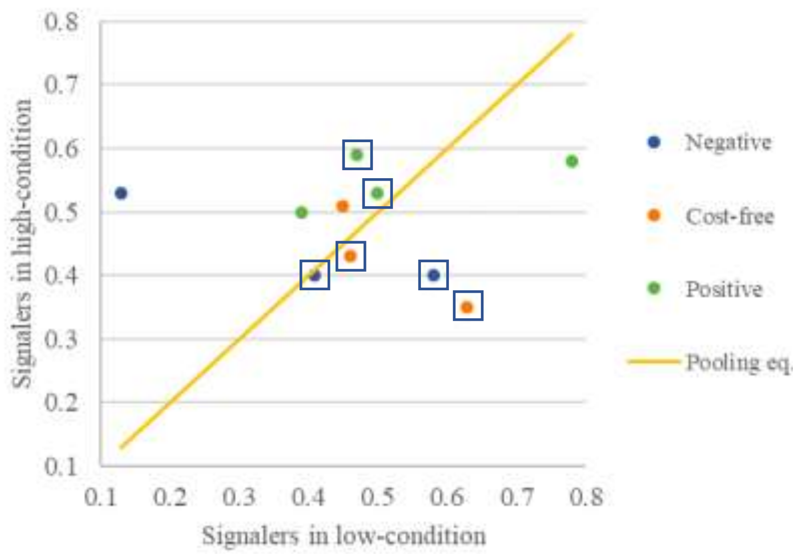


Figure 1 shows the use of the two signals by the state of the signaller in ten experimental groups in the no trade-off treatment. X axis shows how often one of the two signals was chosen by low-quality players. Y axis shows the same for high-quality players. Yellow line represents the equiprobable use of signals by the two types. Please note that in treatment without trade-off both signals can function as a low or high intensity signal, therefore, we coded low and high intensity signals based on the direction of the potential separation (the divergence from this line). We coded low and high-intensity signals in one way for groups above the yellow line and for groups under the yellow line we used reverse coding. Groups that played after a game where the expected outcome was an honest equilibrium are presented in squares.

H2b

In the trade-off condition we do not expect honest signalling (H2b), because for high-quality signallers the marginal benefit resulting from the use of high-intensity signal (e.g. in negative equilibrium cost condition: +100+1200 = 1300) is not higher than the marginal cost of giving high instead of low-intensity signal (ibid.: 1600-100=1500). The only explanation for participants' aspiration towards separating equilibrium could be that individuals played half of the games as a receiver and thus they could have tried to increase their chances of success as a receiver by paying the cost of honesty in the role of high-quality signaller (which role was played only in 25 percent of the game as opposed to 50 percent as receiver). The probability that high quality signallers in high-quality select the low intensity signal is lower (β=-1.787, p<0.001, Model 2a, Table 17) but this separation does not intensify over time in the trade-off condition (β=-0.056, p>0.05 see Signaller's state-round interaction effect in Model 2b, Table 17). We find significant differences between equilibrium cost treatments. The probability of selecting low-intensity signals is lower among low-quality signallers in the negative cost treatment compared to the other two equilibrium cost treatments (β=-1.304 p<0.01 Model 2c, Table 22) and high-quality signallers are less likely to use to high-intensity signals in the negative cost treatment (see Signaler's state – equilibrium cost interaction, Model 2c, Table 22).

H2c

The conditional trade-off manipulations clearly influenced which signal was sent by different types of signallers. High-quality signallers used low-intensity signals with a lower probability (or in other words high-intensity signals with a higher probability) than low-quality signaller ($\beta = -3.386$, $p<0.001$ see the effect of Signaller’s state in Model 3a, Table 17). Moreover, the use of low-intensity signals by high quality signallers decreased as time passed ($\beta = -0.152$, $p<0.001$ see Signaller’s state – round interaction effect in Model 3b, Table 17), which means a gradual progress towards a separating equilibrium. In the negative cost treatment, the probability of selecting low-intensity signals is lower among low-quality signallers at the beginning ($\beta=-1.574$, $p<0.05$ Model 3c, Table 22) but they are more likely send low-intensity signals honest to a higher extent over time compared to the other two equilibrium cost treatments ($\beta=0.170$, $p<0.05$ Model 3c, Table 22). As time passes, high-quality signallers switch to high-intensity signals to a higher extent in the negative equilibrium cost treatment than the other two equilibrium cost treatment ($\beta=-0.202$, $p<0.01$ Model 3c, Table 22).

Low-intensity signals were more likely when the “O vs. //” signs were used ($\beta=0.573$, $p<0.01$, Model 3a, Table 17) in comparison to other signal pairs (“<> vs. []” or “~ vs.)(”).

The probability of refusing resources was also higher when ”O vs. //” signs were used ($\beta=0.400$, $p<0.05$ Model 3a, Table 20), probably because of the increased use of low-intensity signals (”O”) mentioned previously.

Section 5. Instructions

INSTRUCTIONS

The experiment will consist of 60 rounds of decision-making, each round is one game. In each round, you will be paired with other participants. All participants are going to play the game simultaneously.

Before each round, the computer randomly pairs two players who will play together in that round. One member of the pair will be player X and the other will be player Y. In each round, these roles are also randomly assigned by the computer.

Moreover, there are two types of Y in the game: half of the Y players will be in a BLUE state and the other half will be YELLOW. These states are also randomly drawn by the computer at the beginning of the game.

Players know that they are player X or Y. But only Y knows that she/he is BLUE or YELLOW. Y's status is unknown to X player. In each round, 20 players are divided into 10 pairs. The table below shows a possible allocation of 10 pairs:

20 participants																			
Pair 1		Pair 2		Pair 3		Pair 4		Pair 5		Pair 6		Pair 7		Pair 8		Pair 9		Pair 10	
X	Y (b)	X	Y (y)	X	Y (b)	X	Y (y)	X	Y (b)	X	Y (y)	X	Y (b)	X	Y (y)	X	Y (b)	X	Y (y)

*(b) – BLUE, (y) – YELLOW

The game is the following: X owns a resource, and Y wants to obtain that resource. If Y succeeds to request the resource from the player X - whether Y is the BLUE or the YELLOW - it means HUF 1200 to Y. If the player Y fails to get the money, she/he gets HUF 0. The possible payoffs of Y are shown in the following table:

Y’s payoffs along with different decisions:

	X’s decision	
Y state:	Give the resource	Do not give the resource
BLUE	HUF 1200	HUF 0
YELLOW	HUF 1200	HUF 0

X does not know the state of Y. However, X wins only in those cases if she/he gives the resource to Y who is in the BLUE state or does NOT give the resources to those Ys who are in the YELLOW state. The possible payoffs of X are shown in the following table. If player X gives the resource to player Y in the BLUE state, X wins HUF 1200, but if X gives the resource to player Y in the YELLOW state, X does not get anything. Player X also earns HUF 1200 if X does NOT give the resource to the YELLOW players, and conversely, if player X does not give the resource to Ys in the BLUE state X does not win anything.

X’s payoffs along with different decisions:

	X’s decision	
Y state:	**Give the resource**	**Do not give the resource**
BLUE	HUF 1200	HUF 0
YELLOW	HUF 0	HUF 1200

Y player can use two signals to request the resource from X:

<> (peak brackets) and [] (square brackets).

These marks may reduce, increase, or make no changes on the payoff of player Y in the game. The way it affects the payoffs will appear next to the sign during the game.

Example 1.

Y can select from the following signs: (1) using the <> sign reduces Y's payoff by HUF 300; (2) using the [] sign increases Y's payoff by HUF 300. Y decided to use the <> sign and X gave him the resource. In this case, the payoff of Y will be 1200 - 300 = 900.

Example 2.

Y can select from the following signs: (1) using the <> sign has no effect on Y's payoff; (2) using the [] sign reduces Y's payoff by HUF 500. Y decided to use the <> sign and X gave him the resource. In this case, the payoff of Y will be 1200 - 0 = 1200.

Example 3.

Y can select from the following signs: (1) using the <> sign reduces Y's payoff by HUF 300; (2) using the [] sign increases Y's payoff by HUF 300. Y decided to use the [] sign and X gave him the resource. In this case, the payoff of Y will be: 1200 + 300 = 1500.

Example 4.

Y can select from the following signs: (1) using the <> sign reduces Y's payoff by HUF 300; (2) using the [] sign increases Y's payoff by HUF 300. Y decided to use the <> sign and X did NOT give him the resource. In this case, the payoff of Y will be 0 - 300 = -300.

Each round is important since, at the end of the game, participants will be paid on the basis of two randomly selected rounds where they played role X and Y. If a participant's payoff is negative, then we rounded the amount up to HUF 0. In addition to this calculated payoff, participants receive a show-up fee of HUF 1000.

The time available for your choice will display in the top right corner of the screen.

Thank you for your participation!

Have fun and good luck!

FURTHER INSTRUCTION I.

In the next game, Player Y can use one of the following signals depending on their colour to get the resource from player X. Under the signals you can see how each signal can change your earnings.

< >	**[ ]**
The use of this signal in this game **reduces** the earnings of Player Y in a **BLUE** state by **HUF 100**	The use of this signal in this game **increases** the earnings of Player Y in a **BLUE** state by **HUF 100**
The use of this signal in this game **reduces** the earnings of Player Y in a **YELLOW** state by **HUF 1500**	The use of this signal in this game **does not change** the earnings of Player Y in a **YELLOW** state

S5 Screens of the experiment for Player Y

Which round out of 60: 2 Remaining time [sec]: 16

State

In this round the computer has classified you as “Player Y”. Your randomly generated state is:

YELLOW

Please choose one of the signals that can be sent to Player X:

- [] The use of this signal in this game decreases your earnings by HUF 100.
- <> The use of this signal in this game increases your earnings by HUF 100.

Which round out of 60: 2

Please wait until Player X makes a decision.

Next

.

Which round out of 60: 2

In this round you were Player Y.
Your randomly generated state was:
YELLOW

The signal you used:
<>

The decision of Player X:
Gave the resource.

Your decision was successful.

Your earning from the resource is: HUF 1200
The use of this signal in this game changes your earnings by: HUF 100.
Your total earnings in this round: HUF 1200

Next

.

Screens of the experiment for Player X

Which round out of 60: 2

In this round you were Player X.
Please wait until Player Y makes
a decision.

Next

.

Which round out of 60: 2 Remaining time [sec]: 11

Signal

Player Y sent you the following signal:

<>

Please choose one of the following options:

- I give him/her the resource.
- I do not give him/her the resource.

Next

Which round out of 60: 2

The randomly generated state of Player Y was:	The signal from Player Y was:	In this round you were Player X. Your decision was:
YELLOW	<>	**You gave the resource.**

Your decision was unsuccessful.

Your earning from the resource is: HUF 0

Next

.

5. SOCIAL PREFERENCES AND THE EVOLUTION OF FRIENDSHIP AND NEGATIVE RELATIONS IN PRIMARY SCHOOLS [11]

5.1. Introduction

Understanding the determinants of the existence and perpetuation of cooperative behavior is the subject of intensely multidisciplinary research agenda (Haldane 1955, Hamilton 1963, Hamilton 1964, Smith 1964, Singer and Fehr 2005, Hooker et al. 2008, Rand et al. 2012, Nowak 2006, Dufwenberg and Kirchsteiger 2004, Yamagishi et al. 2013). Social relationships and their dynamic nature help explaining individual social preferences, i.e., the willingness of individuals to endure personal sacrifices that help overcome social dilemmas (Fehl et al. 2011, Rand et al. 2011, Bravo et al. 2012, Wang et al. 2012, Melamed et al. 2017, Melamed et al. 2018, Takács et al. 2021. The fact that some individuals are acting more prosocially than others could be attributed to a variety of psychological, economic, cognitive, and social determinants (Thielmann et al. 2020, Singer and Fehr 2005, Rand et al. 2012, Nowak 2006), but could also be reasoned by the social benefits of acting or appearing to be pro-social. The key social driver of prosocial behaviour is arguably the expectation that favours will be reciprocated (Binzel and Fehr 2013b, Pletzer et al. 2018). Reciprocation can manifest in various forms (Willer 2009), including the creation, disruption and rewiring of social ties after fruitful or exploitative collaboration (Rand et al. 2011). Rewarding and punishing via network changes, labelled as link reciprocity (Rand et al. 2011), has been detected under controlled experimental conditions (Rand et al. 2011, Fehl et al. 2011, Wang et al. 2012, Melamed et al. 2017, Melamed et al. 2018). The direct maintenance or potential termination of existing partnerships in repeated interactions can facilitate prosocial and deter selfish behaviour (Fehl et al. 2011, Rand et al. 2011, Melamed

[11] I am grateful to Simone Piras and Simone Righi for the data that they made available for me, and for the comments that they gave me on this chapter.

et al. 2018) via the desire to maintain beneficial social ties (Herman-Stahl and Petersen 1996, Bicchieri 2005, Bond et al. 2007, Ueno 2005, Knifsend et al. 2018). Among members who did not form a partnership before, indirect link reciprocity - such as proposing new partnerships to prosocial individuals and avoiding individuals with bad reputation also plays a role in maintaining prosocial behaviour (Fehl et al. 2011, Rand et al. 2011, Wang et al. 2012, Melamed et al. 2017, Melamed et al. 2018) if social ties are beneficial.

Link reciprocity is expected to have observable consequences on the social network. First, the preference for creating or maintaining relations with prosocial partners should be reflected in the degree distribution in a partnership network (Perc and Szolnoki 2010, Santos et al.2006, Santos et al. 2008). Second, if a tie can be terminated unilaterally (Hauk and Nagel 2001, Coricelli, D. Fehr, and Fellner 2004, Wang et al. 2012, Vörös et al. 2019), partnerships between prosocials are more likely to survive. This can be translated into a network pattern, where similarly prosocial individuals are sorted into partnerships (Takács et al. 2008, Wang et al. 2012, Melamed et al. 2018). While experimental research provides strong evidence of generosity-based network dynamics (Fehl et al. 2011, Rand et al. 2011, Wang et al. 2012, Melamed et al. 2017, Melamed et al. 2018), empirical research conducted on real-life networks gives ambiguous results (Brañas-Garza et al. 2010, Kovářík et al. 2012, Logis et al. 2013, Dijkstra and Berger 2018, Shin et al. 2019). In the following, we summarize the results of field studies that specify friendship as a partnership that can be affected by prosociality-based mechanisms, because it is associated with support, trust, and kindness (Vörös and Snijders 2017). These studies were usually undertaken in a school environment, with a few exceptions (Binzel and Fehr 2013a). Indeed, school environment offer the possibility of collecting complete network data, which is extremely valuable when assessing indirect link reciprocity. Empirically studying the social network determinants of prosociality implies the need for accurate relational data referred to the subject of study. The school setting, with stable social groups that repeatedly and systematically interact with each other provides the researcher with the ideal setting to collect relational data (Molano et al. 2013, Logis et al. 2013, Dijkstra and Berger 2018, Shin et al. 2019). Furthermore, studying the social preferences of children opens a window to the corresponding behaviors in adults. Indeed, social preferences are acquired during childhood and tend to remain stable once the child becomes an adolescence (Caprara et al. 2000, Eisenberg et al. 1999) and we can further surmise that they influence attitudes in adulthood.

Evidence from field studies is weak regarding the first proposed consequence of prosociality-based partner selection, namely the higher embeddedness of prosocial individuals. The simple calculation of how many friends an individual has, shows clear correlation with prosociality in economic games (Kovářík et al. 2012). In these studies, however, participants in the games come from the same pool as friends, therefore it is difficult to distinguish between two mechanisms: the preference for prosocial friends and the fact that donations to friends are higher (Leider et al. 2009, Goeree et al. 2010, Brañas-Garza et al. 2010, Binzel and Fehr 2013a). Indeed, even if the games are anonymous and the matching is random, a decision maker with more friends in the classroom is more likely to be partnered with a friend. Previous studies tried to distinguish these two competing explanations by exploiting the fact that friendship ties are asymmetric (Vörös et al. 2019) and incoming and outgoing friendship nominations can be analysed separately. Potentially, ties leading to individuals mirror partner selection and higher outgoing friendship nominations reflect the association between donation levels and the number of friends an individual has in the classroom. Upon closer examination, in cross-sectional studies friendship indegree (the number of received nominations) shows no relation with prosocial behaviour (Brañas-Garza et al. 2010, Kovářík et al. 2012), while the number of nominations made (outdegree) does (Kovářík et al. 2012), which may indicate the absence of partner selection and the presence of higher generosity towards friends. Contradicting findings are provided by studies using Social Network Analysis (SNA), which could also estimate the tendency of sending and receiving friendship ties separately. With a few exceptions (Logis et al. 2013), social network analyses reveal that prosocial individuals tend to receive more friendship nominations from classmates (Shin et al. 2019, Dijkstra and Berger 2018). In this paper we retest the theoretical prediction about the preference for prosocial partners among young children. We hypothesize that children are more likely to nominate a prosocial classmate as friend:

H1 Prosocial children are more likely to be nominated as good friends.

Scientific evidence is mixed for generosity-based partner selection as well. The assumption that prosocial individuals tend to befriend with each other has been proven both in cross-sectional (Leider et al. 2009) and in longitudinal studies (Logis et al. 2013, Shin et al. 2019). Other studies, however, do not report significant results on assortative friendship formation (Binzel and Fehr 2013a, Molano et al. 2013, Dijkstra and Berger 2018). We contribute to this debate by testing two hypotheses about assortative mixing by social preferences. Our first assumption is based on the observation that both selfish and prosocial individuals have the same number of friends (Leider et al. 2009) but as a result of sorting, there has to be a quality difference in their friendships (prosocials befriend other prosocials and selfish students have selfish friends). Accordingly, we test whether generosity-based sorting affects both prosocials and selfish individuals.

H2a Children are more likely to create or maintain friendship relations with similarly prosocial others.

It could be the case, however, that friendship is formed only by prosocial individuals, and selfish group members are excluded or remain on the periphery of the network (Takács et al. 2008). Therefore, we hypothesize and test whether mainly prosocial children befriended each other:

H2b Prosocial children are more likely to create or maintain friendship relations with similarly prosocial others.

By mapping only positive affections, such as friendship, previous research fails to capture repulsive forces that also shape social networks in a real-world setting (Stadtfeld et al. 2020). Negative relations are particularly important in the context of prosocial behaviour. It has been showed that selfish behaviour evokes negative emotions and a desire to punish selfish individuals (Fehr and Gächter 2002, Fehr and Fischbacher 2004). In light of this, we formulate a hypothesis about the higher probability of negative ties towards selfish individuals.

H3 Selfish children are more likely to receive or maintain negative ties.

If punishment involves a cost (e.g., retaliation Denant-Boemont et al. 2007, Nikiforakis 2008), a second-order social dilemma arises (Henrich et al. 2006). The dilemma is about undergoing the cost of punishment for the benefit of other group members because punishment can deter selfish behaviour (Boyd and Richerson 1992). Previous studies show that prosocial individuals are more likely to engage in punishment of non-cooperators (Fehr and Gächter 2000, Fehr and Gächter 2002, De Quervain et al. 2004). We thus generate a hypothesis on the tendency of prosocial individuals to send negative ties towards selfish individuals.

H4a Prosocial children are more likely to create or maintain negative relations to selfish others.

Opposite arguments can also be formulated. Research on antisocial punishment shows that selfish individuals tend to punish individuals with higher prosociality (Herrmann, Thoni, and Gächter 2008, Gächter and Herrmann 2009, Brañas-Garza et al. 2014), which could be motivated by revenge (Nikiforakis 2008), competition within (Zizzo 2003, Fliessbach et al. 2007, Samu et al. 2020) or between groups (Gächter and Herrmann 2009, Sylwester et al. 2013). If punishment does not go just in one direction, from prosocials to proselfs, but from proselfs to prosocials as well, we expect generosity-based repulsion:

H4b Negative relations are more likely to occur between children with different social preferences.

To contribute to this mixed scientific evidence, the aim of this study is to retest the role of social preferences in partner selection in schools. For this purpose, we combine the strengths of past research that advanced behavioural measures and network analysis. We employ network evolution models to assess how friendship formation is affected by individuals' social preferences elicited through three incentivised tasks. Besides the methodological innovations, our dataset also comprise data on negative relationships, which allows us to shed a unique light on the role of social preferences in their formation. Further, we are able to document processes that segregate individuals with different social preferences.

5.2. Data and method

5.2.1. Participants

Data were collected in nine primary schools in northern Italy among students in grades 4 and 5, with a median age of 10 years. A total of 420 students from 20 classes took part in at least one step of the research, but only 405 are retained for the analysis (49.1% women). We sought parental consent and we received it for all children of all the classes but one, who, for this reason, did not participate in the study. We surveyed the social networks of the classes in two waves, at the end of 2017 and in May 2018. Students' social preferences were measured in the first wave through incentivised economic games. Data collection for the games was performed on tablets or netbooks provided by the Reggio Emilia Behavioral and Experimental Laboratory (REBEL). The games were programmed using the web-based platform oTree (Chen et al. 2016). The questionnaires were filled on paper with the support of the research team and the teachers.

5.2.2. Measurements

(a) Social preferences

We use incentivised economic games with the aim to elicit real preferences rather than measuring theoretical statements and attitudes. During the games, apart from a pen of the University of Bologna as a show up fee, children were collecting tokens, and depending on the number of tokens collected, they were rewarded with zero to two cinema tickets. The computer ranked children in each class. Kids in the top tercile received two cinema tickets, those in the second tercile one. The children played four tasks but the payoffs for each of them, as well as the total one, were displayed only at the end of the full session. On the last oTree page, students saw how they had performed in the games and to which third of the rankings they belonged to. Rewards were handed over privately and we also protected the anonymity of children by giving them identification numbers.

Among the many decision games that have been designed to measure social preferences, we chose to implement three standard ones, with the objective to keep the design simple and accessible to children of 10 years. Between each of the games, groups were reshuffled (stranger's treatment), and in all cases groups where anonymous, with students knowing only they would play with someone else in their class. The first game, the Public Goods Game (PGG) Ledyard and Palfrey 1994 was played in groups of four. Each received 40 tokens and then decided how much to contribute to a group account – 0, 10, 20, 30 or 40 tokens – knowing that twice the total contribution would be divided equally among the four participants. This game aims at measuring the individual willingness to contribute to the creation of a public good that benefits every component of the social group.

In the second game, Dictator Game (DG) (Kahneman et al. 1986), students received 100 tokens and could donate any discrete amount between 0 and 100 to a classmate with whom they were randomly and anonymously paired. Everybody played both the roles of the donor and of the recipient, but at the end of the session the computer selected only half of the players to be the donors, and both their and their peers' payoff was defined by their decision. The dictator game was introduced to study the pure prosociality, i.e., the willingness to pay a cost for the benefit of another individual.

In the third game, Trust Game (Berg et al. 1995), students got again 100 tokens, and had to implement two tasks. Trustors (TG) had to decide how much to give to their partner, knowing that they (TGB) would receive three times the amount donated and could give back any amount from the obtained tokens (including zero). Participants made decisions in both roles and again the computer decided randomly which decisions would be payoff-relevant. Participants were made aware of this process and explained with examples. This game was played after the PGG and the DG, to ensure that the received amount would not affect further decisions unintendedly. Since all participants played these games in the same order, previous decisions certainly influence subsequent choices, however, the absence of feedback between tasks reduces the scope for past decisions' outcomes to influence future decisions. All tasks included 2-4 test questions to assess whether the participants understood the game. All questions were answered, and further doubts raised by the participants addressed before taking the payoff-relevant decisions. Further, all tasks included in the instructions examples and vignettes to enhance understanding.

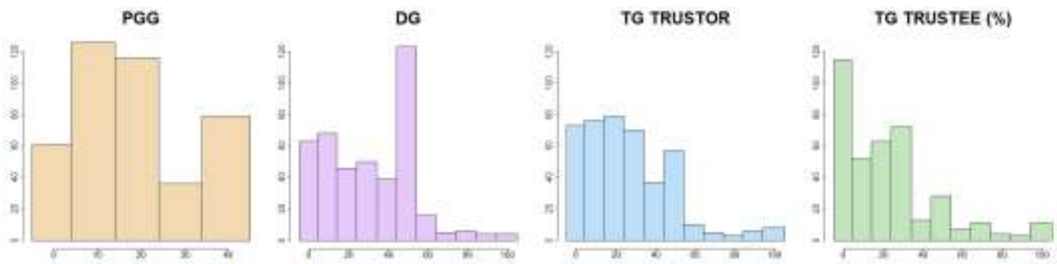


Figure 1. Distribution of allocations in the four economic game decision (PGG, DG, TB, TBG)

(b) Social Relations

A snapshot of a class network contains essential information about the social processes taking place in the classroom. We collected social network data by asking students to evaluate their peers in the classroom along various dimensions in two waves. Longitudinal data allow us to measure how the relationships in the classroom change over time. In the following, we perform two analyses to examine both positive and negative ties as dependent variables. To study how friendship changes, participants were asked to rate how much they consider their classmates as friends. Classmates' names were listed in the questionnaire and next to each name, respondents could select a response on a 5-point scale: 1 - good friend, 2 - friend, 3 - indifferent, 4 - not friend, 5 - absolutely not friend. For the analysis we restrict our attention to good friend nominations. If someone nominated another classmate as good friend, that tie is set to 1, and everything else become 0. Based upon these nominations, we constructed one network for each of the classes, where nodes represent children and links between them represent good friend nominations.

In addition to examining friendship dynamics, we investigate how negative relationships develop. To assess the characteristics of negative relations, we similarly constructed one networks for each class, where links between them represent negative ties. Due to the sensitivity of negative nominations, these are naturally sparser. For this reason, we constructed a measurement combining multiple network items. A negative tie exists if at least one negative nomination was sent in the following three questions. First, from the friendship network, not friend and absolutely not friend nominations were treated as negative ties. Second, students were also asked how nice they think others are on a 5point scale. The two values at the negative end of this scale, namely unpleasant and very unpleasant, were classified as negative. The third question used to build the negative networks asked about the desired desk mates. A negative tie is considered to exist if children nominated a peer as someone, they ‘would not sit next to’, the other options being ‘would sit next to’. In sum, a tie is set to 1 if at least one of five potential nominations (not friend, absolutely not friend, unpleasant, very unpleasant, would not sit next to them) is present, 0 if none of them exist. Given that nominations might not be reciprocated, both the positive and negative networks are directed networks, where the fact that A is connected to B does not imply that B is linked to A.

Figure 2. (a) The overlap between the network questions (b) Friendship (yellow) and negative (blue) networks in one class. Squares represent girls and circle represent boys. Colours of the node indicate social preference measured by the dictator game. Dark purple nodes offered at least half of their tokens, while light purple shows students who did not give any token to their classmates. Layout is defined by the good friend nominations using the Fruchterman-Reingold algorithm.

The negative networks constructed in this way are almost identical to the network defined by the rejection of being desk mate, 84% of the ties in the merged network correspond to such avoidance (see column 5 in Table 1). The other four nominations (unpleasant, very unpleasant, not friend, absolutely not friend) overlaps by only 21-29% with the constructed network (ibid.). The reason why these percentages do not add up to 100% is that there is also overlaps between these nominations (see Figure 2a).

(c) Controls

In both analyses of positive and negative relations, we control for dynamics determined by gender and school performance. We include gender because it has a well-known positive effect on giving in the DG (Croson and Gneezy 2009, Horn et al 2022), and at the same time children prefer to select same-gender friends (McPherson et al. 2001, Stadtfeld et al. 2020). In the data, boys are coded as 0 and girls as 1.

Similarly, school performance may also correlate positively with social preferences

(Caprara 2000) and network dynamics (Lomi et al. 2011). Disregarding these factors could lead to confounding. On the one hand, we might overestimate the preference for prosocial individuals. On the other hand, a higher likelihood of friend formation between well-performing students could be misinterpreted as befriending by prosociality. Here, school performance is measured as the average score in the last tests in Italian language and mathematics, with a range between 0 and 10.

Besides gender and academic achievement, popularity-based homophily was also revealed in previous studies (Haselager et al. 1998, Newcomb and Bagwell 1995). Although popularity is expected to be positively related to prosocial preferences, it has a distinct effect on the development of friendship relations (Logis et al. 2013). If we do not control for the assortative bonding of popular kids, we may overestimate the impact of social preferences, if popularity and prosociality correlate (Logis et al. 2013). In our data, popularity is measured by a network question, where students were asked to nominate those who they think are ‘liked by everyone’ or ’liked by many others’ in the class. For each child, we divided all incoming nominations by the size of the class (excluding children that were absent on the day of the survey) and after a transformation of this ratio to discrete numbers between 0 and 10 we included it as a node attribute into our models.

5.3. Analytical approach

We apply the standard statistical methodology used to study network dynamics in small, closed groups, called Stochastic Actor Oriented Models (SAOMs) (Snijders et al. 2010). This method enables us to test theories about social mechanisms influencing the formation, persistence, and termination of relationships. To estimate SAOMs, the R package RSiena has been developed and constantly maintained by network researchers Ripley et al. 2011. SAOMs model the evolution of social ties over time by simulating mechanisms that may be responsible for changes between the empirically observed snapshots of the networks. At each step of the simulations, an actor - chosen at random - can create, maintain, or terminate a tie as a result of the state of the network in a given step of the simulation, which are sequentially modelled (Ripley et al. 2011). Individual characteristics such as gender or social preferences are used to explain network changes as well. The effects of network structures and individual or dyadic attributes on tie formation are expressed as log odds of tie creation or maintenance.

Several well-known network dynamics can be assessed during the estimation process, such as the general tendency of sending positive or negative social ties (outdegree), their bilateral nature (reciprocity), the closure of triadic relations (transitivity), or even higher-level structural effects. We listed the specifications of the effects in our models in SM Table 1. After controlling for these structural effects, we are able to test the hypothesized effects of prosociality on tie creation or maintenance.

The decisions in our economic games represent node level characteristics that may affect tie creation. Since decisions in the games are correlated, we run separate models for each game. Effects related to our main hypotheses are specified in a similar way for each model, using the following three effect specifications. First, we surmise that prosocial behaviour can influence the tendency of receiving a friendship or a negative tie. The altX effect captures whether individuals consider generous classmates as friends with higher probability (H1) and create or maintain fewer negative ties with them (H3). This measure is defined by the social preference (v) of the receiver (j) of a tie (*x*, an indicator variable which can assume value 1 or 0 depending on whether it exists or not) from the actor (i) who is selected in a given step of the simulation: $\sum_j x_{ij} v_j$.

Second, we projected a dyadic effect based on a common attribute of the sender and the receiver. To assess whether individuals tend to befriend or avoid negative ties with similarly prosocial others (H2a, H4b) we use an effect called simX, which is the sum of centered (for each class) similarity scores between i and j: $\sum_j x_{ij}(sim^v_{ij} - \widehat{sim^v})$, where similarity scores are adjusted to the maximum distance in prosociality $sim^v_{ij} = \frac{\max_{ij}|v_i - v_j| - |v_i - v_j|}{\max_{ij}|v_i - v_j|}$ and $\widehat{sim^v}$ is the mean score. High values indicate nominations between similarly prosocial peers either in the friendship or in the negative network and low values refer to dissimilar relations.

Additional hypotheses states that only prosocial individuals sort themselves into friendships (H2b) and only prosocials bear the cost of negative ties towards selfish classmates (H4a). To test these, we use an interaction effect between the above mentioned *simX* effect and *egoX* which measures the tendency to send more friendship or negative ties by prosocial individuals. Using mathematical terms, *egoX* is the sender's out-degree weighted by sender's prosociality: $v_i x_i$. The interaction of these two effects allows us to assess significant differences in similarity-based tie formation along the sender's prosociality. If prosocials are more likely to nominate similar others in the friendship network, and less likely to do so in the negative network compared to selfish children, the interaction of simXegoX effect will be significant positive and negative, respectively.

Using these effects, we run 2x4 models for changes in the two networks (friendship and negative) based on the four decisions made in the three economic games (PGG, DG, TG, TGB) and we run each model with and without the interaction term between the activity of prosocial students (egoX) and homophily between children with similar social preferences (simX). We transformed all decisions in the games into a reduced scale between 0 and 10 (0 and 4 in PGG) by dividing the original values by 10, because RSiena estimations are scale-dependent (Ripley et al. 2011 p. 27). Separate models with almost the same specification (more detail in Supplementary Material (SM)) are estimated for each of the 20 classes in our sample. It is possible to use different model specification if needed, because in the meta-analysis each effect is tested independently from other effects in the model. Models are adjusted slightly by fixing certain effects to 0 if standard errors were too large. In the SM it can be traced which effects were fixed during the estimation process (see N in SM tables about the results of the meta-analysis). The separate SAOMs for each class are also included in the SM. To summarize the results, we apply meta-analysis which we implement afterwards to the estimates produced by the separate SAOMs. We assume that the collected networks are a sample from the population of networks, and we infer the mean population value for each effect using the restricted maximum-likelihood estimator from the 'metafor' R package (Viechtbauer 2010).

In the following, before we discuss the results of the meta-analysis, we provide descriptive statistics on the decisions in the games, and we give a general overview of friendship and negative networks.

	1. Jaccard coeff.	2. Reciprocity	3. Outdegree	4. St. outdegree	5. Overlap with the constructed negative network
good friend	0.46	0.69	5.74	0.27	0.00
unpleasant	0.12	0.88	1.31	0.06	0.29
very unpleasant	0.18	0.92	0.99	0.05	0.21
not friend	0.11	0.89	1.24	0.06	0.26
absolutely not friend	0.24	0.92	1.05	0.05	0.22
would not sit	0.35	0.74	3.90	0.18	0.84
constructed negative network	0.38	0.72	5.02	0.23	1.00

Table 1. Network descriptives

5.4. Results

5.4.1. Descriptive results

(a) Social preferences

In the PGG, the average contribution was 18.76 tokens, with a standard deviation of 13.18. 14.53% contributed with 0 tokens (Figure 1). In the DG, 32.45% of the children donated at least half the amount and 9.93% gave nothing to the other player (ibid.). The average offer was 30.53 tokens, with a standard deviation of 21.71. In the first step of the TG, 25.76 tokens were sent to trustees on average (SD = 21.52). 46 students (11.36%) did not receive anything and therefore they could not decide in the second step, where children returned an average of 22.30% of their available amount to the trustor (SD = 23.27). 19.57% did not return anything (ibid.).

Correlation is strong between DG and TG decisions ($\rho_{DG,TG}$ =0.61,$\rho_{DG,TG}$ = 0.43,$\rho_{TG,TGB}$ = 0.47,$p < 0.001$), and we only see moderate correlation between the PGG and the other two games ($\rho_{PGG,DG}$ = 0.34,$\rho_{PGG,TG}$ = 0.28,$\rho_{PGG,TGB}$ = 0.19,$p < 0.001$).

(b) Friendship ties

The network of good friends contains 2495 ties in the first wave and 2128 in the second, and while these numbers are close, only half of the nominations remained stable over time. In Table 1 the Jaccard coefficient shows the proportion of stable ties compared to the union of the ties in the two waves (t11/(t01 + t10 + t11)). Overall, 69% of the good friend nominations are reciprocated (column 2, ibid.). An average of 5.74 individuals are nominated as a good friend, so children typically have six good friends (column 3, ibid.). In our sample, 27% of all the possible ties between two children is a good friend tie, taking into account absent children who could not nominate anyone (column 2, ibid.).

(c) Negative ties

While friendship ties decreased between the two waves, the total number of negative ties – sent by either of the nodes – increased from 1935 to 2323. Only 38% of these ties are stable over time (see column 1 in Table 1). Overall, 72% of the negative ties are reciprocated (column 2, ibid.). On average, one child has 5.02 negative relationships (column 3, ibid.), and a quarter of all possible pairs in the class have a negative relationship (column 4, ibid.).

5.4.2. Results of the meta-analysis

We summarize the results of separate SAOMs using a meta-analysis. Results of the meta-analysis are illustrated in Table 2 for friendship ties and in Table 3 for the negative network. Detailed tables in the SM provide information about all the effects included in the models and the complete results for the meta-analysis.

Our first hypothesis hypothesises preference for prosocial individuals as friends (H1). More precisely, we tested whether a student with higher contribution in the four economic games is more likely to be nominated as a good friend by their peers. The estimated altX effects and the corresponding p-values in Model 1 in Table 2 demonstrate a lack of evidence for generosity-based partner selection. The PGG altX parameter does not differ significantly from zero. Results are the same for the other games: the null hypothesis, namely that the effect is null in all groups, cannot be rejected.

We do not find evidence for the second hypotheses either, in which we expected assortative partner matching between similar peers (H2a) or protective friendship formation of prosocial individuals (H2b). Neither simX (Model 1 in Table 2) nor the interaction between egoX and simX are significantly different from zero (Model 2 in Table 2).

Next to these non-significant results, robust effects are found for well-known structural effects (e.g., reciprocity, transitivity, outdegree - popularity) and for the control variables, including the preference for same sex friends, and the preference for popular classmates. We also demonstrate that popular individuals are more likely to be nominated as good friends (see extended table of the results of the meta-analysis in the SM).

	Friendship									
	Model 1					Model 2				
effect	fig	est	se	p	Q_p	fig	est	se	p	Q_p
altX PGG *(H1)*		0.023	0.028	0.398	0.560		0.018	0.036	0.623	0.465
simX PGG *(H2a)*		-0.009	0.116	0.935	0.082		-0.018	0.111	0.872	0.150
egoX x simX PGG *(H2b)*							0.049	0.101	0.629	0.546
altX DG *(H1)*		-0.003	0.019	0.885	0.503		0.034	0.031	0.284	0.528
simX DG *(H2a)*		-0.139	0.085	0.102	0.839		-0.139	0.091	0.127	0.486
egoX x simX DG *(H2b)*							-0.129	0.078	0.098	0.251
altX TG *(H1)*		-0.024	0.024	0.311	0.016		0.012	0.024	0.606	0.359
simX TG *(H2a)*		-0.081	0.121	0.505	0.470		0.009	0.112	0.933	0.718
egoX x simX TG *(H2b)*							-0.129	0.067	0.054	0.870
altX TGB *(H1)*		0.003	0.019	0.865	0.837		-0.003	0.024	0.911	0.916
simX TGB *(H2a)*		-0.009	0.190	0.961	0.213		-0.130	0.215	0.546	0.378
egoX x simX TGB *(H2b)*							-0.092	0.105	0.379	0.083
N=20										

Table 2. Main results of the meta-analysis, Friendship network

The negative network is constructed in a way that measures avoidance (would not sit) to a greater extent than strong emotions (very unpleasant, absolutely not friend) (see overlaps in Table 1). The meta-analysis of the modelled negative network leads to non-significant results with regard to social preferences of the receiver of a tie (H3). For instance, the lower contribution in the PGG does not increase the likelihood of receiving negative nominations within the class (see altX PGG in Model 1 Table 3). The probability of incoming negative ties does not differ significantly between children with generous or selfish offers in the DG (altX DG ibid.). The same applies to the two decisions in the TG (altX TG, altX TGB, ibid.). We find evidence in support of our last hypothesis about the similarity effect in the negative network (H4). We test two forms of this dyadic effect. First, we hypothesized that only prosocial individuals show higher willingness to punish selfish peers via the maintenance of negative ties. We see a contradicting result in the model, where we test the probability of sending or maintaining negative ties by children with contribution in the PGG towards peers who have also high contribution. Here the significant positive effect shows that prosocial children are more likely to nominate other prosocial peers in the negative network ($\hat{\mu}_{PGG} = 0.238$, $p < 0.05$). The interaction of the sender's offers and the similarity between the sender and the receiver of a negative tie in the other three models fail to repeat this result (see egoX x simX effects in Model 2). The heterogeneity test suggests that this effect is homogeneous in the population of networks ($Qp > 0.05$), which strengthens the validity of the result.

	Negative									
	Model 1					Model 2				
effect	fig	est	se	p	Q_p	fig	est	se	p	Q_p
altX PGG *(H3)*		0.002	0.025	0.944	0.493		-0.037	0.031	0.224	0.927
simX PGG *(H4b)*		-0.049	0.110	0.655	0.193		-0.096	0.113	0.395	0.158
egoX x simX PGG *(H4a)*							0.238	0.095	0.012	0.362
altX DG *(H3)*		-0.007	0.017	0.688	0.325		-0.007	0.025	0.768	0.835
simX DG *(H4b)*		-0.201	0.083	0.015	0.118		-0.204	0.086	0.018	0.069
egoX x simX DG *(H4a)*							0.018	0.083	0.828	0.093
altX TG *(H3)*		-0.01	0.02	0.71	0.02		0.02	0.02	0.48	0.40
simX TG *(H4b)*		-0.16	0.11	0.13	0.30		-0.13	0.12	0.27	0.40
egoX x simX TG *(H4a)*							-0.09	0.06	0.14	0.38
altX TGB *(H3)*		-0.00	0.02	0.84	0.89		-0.01	0.02	0.43	0.84
simX TGB *(H4b)*		-0.21	0.21	0.34	0.03		-0.30	0.22	0.17	0.05
egoX x simX TGB *(H4a)*							0.04	0.07	0.58	0.41
N=20										

Table 3. Main results of the meta-analysis, Negative network

We found, however, a significant similarity effect along the DG offers (simX DG, Model 1 in Table 3), which does not depend on the generosity of the sender (see non-significant effect of egoX x simX interaction in Model 2). This means that the probability of having negative relationships Table 3: Main results of the meta-analysis for negative ties decreases as the similarity in individuals' DG offer increases ($\hat{\mu}_{DG} = -0.201$, $p < 0.05$). The test for heterogeneity does not show significant heterogeneity of this effect ($Q_p > 0.05$). This similarity effect has not been found in the TG (see simX TG, simX TGB in Model 1 in Table 3).

5.4.3. Goodness of fit

We give an overview on how well our models perform. The goodness of fit (GOF) of SAOMs can be assessed by contrasting the empirical networks to simulated ones using the parameters estimations from the models. Global characteristics are the bases for comparison that are not fitted by the parameters in the model unequivocally, thus, they can provide information on how well a model represents the empirical data. We examine, for instance, the distribution of indegree, outdegree, geodesic distance (i.e., the length of the shortest path between two children), and triad census (the occurrence of triadic configurations). Dissimilarity between the simulated and the empirical data is evaluated using p-values from a Mahalanobis' distance measure, where a low number indicates a poor fit (Ripley et al. 2011 p. 59). The p-values of GOF statistics are reported in the SM for each model. In the following, we discuss these four indicators in those classes where the model fit was poorer.

Changes in the friendship network are better modelled than the negative network, because fewer classes have p-values under 0.05, where a good fit is rejected. First, we evaluate indegree distributions, where we check whether the observed frequencies by the number of incoming nominations fit well to the simulated values. Indegree distribution fails to fit in the negative network in two classes. The number of individuals without any friendship nomination are close to the confidence intervals of the simulated values in class 7 and individuals with a maximum indegree of five are higher than the upper quartile of the simulated values in class 18 (see related figures in the SM).

The distribution of the children is not increasing gradually along the number of outgoing ties in class 11 in the friendship models. The cumulative number of children by outgoing friendship nominations in the second wave shows individuals without any friendship nomination are higher than expected. We overestimate the number of students with low outdegrees, and since there is a big jump in the empirical network between the outdegree of four and six, we underestimate outdegrees afterwards. The very small p-value suggests that the observed data are far from the simulated data, therefore, we exclude this class from the friendship analysis. Our models for negative ties simulate slightly different cumulative outdegree distributions than it is observed in class 6 and 10, but these classes are not excluded from the meta-analysis.

The third measure (triad census) tests the occurrence of 16 triadic configurations. None of the observed values fit badly with the simulated values calculated by our model on friendship and negative dynamics.

The fourth property examined is geodesic distance. In two classes (6 and 13), the model on negative ties does not fit two classes (class 6 and 13) well regarding the cumulative number of pairs of actors with the shortest path of one, two and three. Due to the small class sizes, the fit of pairs at distance three are poor.

5.5. Conclusion

In this study, we examined the role of social preferences in shaping social networks in a primary school environment. We developed hypotheses along a theoretical framework that emphasizes the role of social preferences in partner choice. Considering friendship and negative relations within the classroom, we tested assumptions about prosociality-based tie formation using the toolkit of SNA. By estimating the dynamics of friendship and negative relations between two waves of data collection, we tested whether prosocial individuals, who have made more generous decisions in economic games in the first wave, receive more friendship and less negative nominations. Unlike other studies (Shin et al. 2019, Dijkstra and Berger 2018), we found no evidence of a preference for more generous friends. Moreover, we fail to prove the existence of repulsive forces towards selfish individuals.

The fact that social preferences are hidden traits, and individuals are not aware of others' social preferences, can explain the lack of impact in terms of networks. Nevertheless, unobserved preferences can be signalled in an observable way. Such signal could be reputational information from others. Reputation and prosocial behaviour are already linked: it has been empirically proven that generosity leads to reputational gains (Macfarlan and Lyle 2015). It is therefore conceivable that the effects of prosociality are mediated by popularity. We see in our models that popularity, indeed, attracts more friendship and less negative nomination, and friendship nominations between popular students are more likely, while negative nominations are less likely (see the results of the meta-analysis in the SM). This reasoning, however, is less valid in our case because, while other studies find a strong correlation between popularity and prosociality (Logis et al. 2013), there is no correlation in this sample ($\rho PGG = -0.189, p < 0.001, \rho DG = -0.013, p = 0.773, \rho TG = -0.046, p =$

0.314, $\rho TGB = -0.011, p = 0.829$).

The lack of evidence of asymmetric relations suggests that class dynamics at this stage of life may no longer be driven by indirect link reciprocity, because students already spent enough time together to gain direct experiences. As opposed to indirect reciprocity, direct reciprocity rather infers symmetric relations. Therefore, we should be able to observe that individuals organize themselves into partnerships with similar peers and maintain negative relations with dissimilar others. Our results do not prove the presence of assortative matching in the friendship network, but we found that dissimilar individuals are more likely to establish or maintain negative relations.

We document that differences in generosity increases the probability of negative relationships among students. Interestingly, we found strong and significant effect only in the DG, whose design is better suited for measuring intrinsic social preferences, while the other games (PGG, first decision in the TG) require more strategic thinking (Thielmann et al. 2020). The second decision of the trust game (TGB) is somewhat closer to the dictator game and shows a similar but non-significant effect. Mutually negative nominations between prosocial and selfish individuals may indicate the combined presence of prosocial and antisocial punishment (Herrmann et al. 2008).

Lastly, a counter-intuitive result has been found. Prosocial individuals are more likely to maintain a negative relationship with each other if we define prosociality based on the contribution in the PGG. Note that this game was played first, and it is more likely that children make more non-strategic decisions in the first game (Anderson et al. 2000). It is also known that intuitive decisions are more cooperative (Rand et al. 2012, Rand et al. 2014, Rand 2016). Along this line, the literature on strong reciprocity also assumes intuitive drivers behind punishment. It might be the case that intuitive individuals establish a negative relationship with each other more easily.

One of the main contributions of this study has been to examine whether social preferences have consequences on the negative networks. Positive networks, on their own, do not represent all the forces that could shape social relationships, and the examination of forces acting in the opposite directions is important (Stadtfeld et al. 2020). One of the limitations of this study is that it does not model these two forces simultaneously. The signs of the studied effects in the two networks (e.g., simX) suggest that these dynamics are not just the two sides of a coin but have distinct effects. Still negative dynamics are underrepresented in research. Hopefully, the present study contributed to fill this gap.

5.6. Supplementary

name	figure	equation	description
density		$\sum_j x_{ij}$	constant, the out-degree if all other effects are zero
recip		$\sum_j x_{ij}x_{ji}$	the tendency to reciprocate a tie
transTrip		$\sum_{j,h} x_{ij}x_{ih}x_{hj}$	the tendency to close open triads in sender's neighbourhood
transRecTrip		$\sum_{j,h} x_{ij}x_{ji}x_{ih}x_{hj}$	the interaction between the transTrip and recip
outAct		x_{i+}^2	sender's activity based on their outgoing ties
inAct		$x_{i+}x_{+i}$	sender's activity based on their incoming ties
inPop		$\sum_j x_{ij}x_{+i}$	the tendency to send a tie based on receiver's incoming ties
outPop		$\sum_j x_{ij}x_{j+}$	the tendency to send a tie based on receiver's outgoing ties
outTrunc		$\min(x_{i+}, p)$	the tendency to be an isolate based on sender's outgoing ties
egoX		$v_i x_{i+}$	sender's activity based on her value in a given covariate
altX		$\sum_j x_{ij}v_j$	the tendency to send a tie based on receiver's value in a given covariate
sameX		$\sum_j x_{ij}I\{v_i = v_j\}$	the tendency to send a tie to receivers with exactly the same value in a given covariate
egoXaltX		$v_i \sum_j x_{ij}v_j$	the tendency to send a tie to receivers with similar value in a given covariate

simX	• •	$\sum_j x_{ij}(sim_{ij}^v - \widehat{sim}^v)$	the tendency to send a tie to receivers with similar value in a given covariate, centered (where $sim_{ij}^v = \frac{\max_{ij}\lvert v_i - v_j \rvert - \lvert v_i - v_j \rvert}{\max_{ij}\lvert v_i - v_j \rvert}$ and $\widehat{sim}^v$ is the mean score)

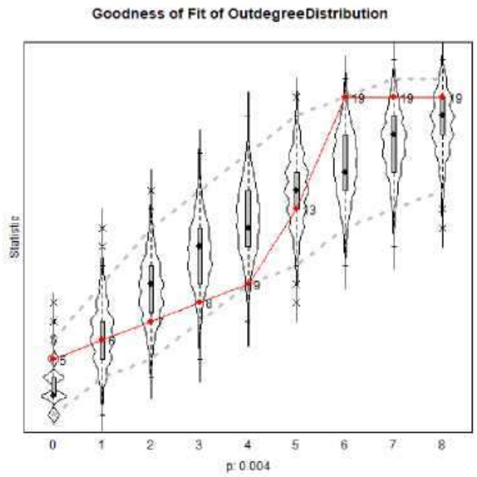
Table 1. See more about the effects in Ripley, R.M., Snijders, T.A., Boda, Z., Vörös, A. and Preciado, P., 2011. Manual for SIENA version 4.0. University of Oxford.

Table 2: Results of the meta-analysis for friendship network and PGG, Model 1

	est	se	N	p	tau2	Q	Qp
outdegree (density)	-1.219	0.2917	20	0.000	0.801	33.791	0.019
reciprocity	1.057	0.1219	20	0.000	0.000	11.539	0.904
transitive triplets	0.174	0.0275	20	0.000	0.001	18.253	0.506
transitive recipr. triplets	-0.043	0.0432	20	0.319	0.064	16.402	0.630
indegree - popularity	-0.024	0.0197	20	0.217	0.000	21.030	0.335
outdegree - popularity	-0.099	0.0204	19	0.000	0.045	22.870	0.196
outdegree - activity	0.025	0.0128	20	0.054	0.034	41.696	0.002
sex alter	0.009	0.0969	20	0.926	0.273	31.901	0.032
sex ego	0.012	0.0753	20	0.872	0.155	23.419	0.219
same sex	0.555	0.0865	20	0.000	0.253	36.779	0.008
grade alter	0.070	0.0430	20	0.103	0.086	22.802	0.246
grade ego	-0.043	0.0414	20	0.299	0.079	23.615	0.211
grade similarity	0.211	0.1492	20	0.158	0.383	30.023	0.052
PGG alter	0.023	0.0276	20	0.398	0.001	17.436	0.560
PGG ego	0.008	0.0385	20	0.828	0.103	33.709	0.020
PGG similarity	-0.009	0.1164	20	0.935	0.239	28.093	0.082
popularity alter	0.185	0.0277	20	0.000	0.048	21.725	0.298
popularity ego	0.013	0.0200	20	0.526	0.026	22.842	0.244
popularity ego x popularity alter	0.024	0.0072	20	0.001	0.012	25.770	0.137

Table 3: Goodness of fit statistics, Friendship and PGG, Model 1

	Indegree distribution	Outdegree distribution	Geodesic Distance	Triad Census
1	0.746	0.535	0.969	0.992
2	0.908	0.095	0.572	0.982
3	0.911	0.641	0.243	0.529
4	0.797	0.428	0.884	0.998
5	0.479	0.070	0.290	0.160
6	0.673	0.773	0.359	0.212
7	0.217	0.698	0.686	0.840
8	0.470	0.707	0.983	0.967
9	0.906	0.933	0.582	0.951
10	0.804	0.623	0.441	0.990
11	0.541	0.004	0.332	0.730
12	0.246	0.726	0.774	0.826
13	0.291	0.271	0.156	0.693
14	0.870	0.786	0.953	0.881
15	0.807	0.982	0.880	0.849
16	0.872	0.442	0.438	0.226
17	0.635	0.804	0.988	0.994
18	0.993	0.531	0.904	0.967
19	0.984	0.694	0.344	0.716
20	0.820	0.712	0.959	0.821



(a) Class 11

Figure 1: Classes with poor fit, Friendship and PGG, Model

Table 4: Results of separate SAOMs for friendship network and PGG, Model 1

	1	2	3	4	5	6	7	8	9	10
outdegree (density)	-0.875 [0.608]	0.29 [1.844]	-5.615 [1.885]	1.178 [1.38]	-2.769 [0.68]	-0.253 [0.815]	-0.017 [0.961]	1.474 [1.154]	-1.314 [0.636]	-1.65 [0.976]
reciprocity	1.151 [0.384]	1.227 [1.063]	1.809 [0.909]	0.588 [0.61]	1.777 [0.537]	0.505 [0.429]	0.671 [0.641]	1.107 [0.403]	1.085 [0.366]	1.889 [0.726]
transitive triplets	0.213 [0.092]	0.45 [0.295]	0.195 [0.211]	0.029 [0.169]	-0.012 [0.098]	0.111 [0.126]	0.279 [0.154]	0.317 [0.124]	0.19 [0.082]	0.312 [0.133]
transitive recipr. triplets	0.048 [0.136]	-0.248 [0.388]	-0.276 [0.284]	0.213 [0.22]	0.194 [0.151]	0.031 [0.185]	-0.062 [0.246]	-0.119 [0.184]	-0.114 [0.137]	-0.251 [0.207]
indegree - popularity	0.018 [0.042]	-0.11 [0.171]	0.12 [0.139]	-0.087 [0.127]	0.049 [0.07]	-0.103 [0.092]	-0.12 [0.102]	-0.332 [0.14]	-0.064 [0.062]	-0.107 [0.094]
outdegree - popularity	-0.207 [0.071]	-0.205 [0.236]	0.048 [0.099]	-0.134 [0.162]	-0.137 [0.05]	-0.087 [0.055]	-0.131 [0.106]	-0.283 [0.083]	-0.097 [0.038]	-0.14 [0.051]
outdegree - activity	-0.041 [0.034]	-0.081 [0.114]	0.245 [0.12]	-0.086 [0.087]	0.093 [0.027]	0.022 [0.032]	-0.093 [0.063]	0.003 [0.034]	0.026 [0.023]	0.067 [0.035]
sex alter	-0.035 [0.189]	0.169 [0.499]	-0.442 [0.57]	-0.601 [0.351]	0.355 [0.218]	0.226 [0.265]	-0.188 [0.362]	-1.062 [0.353]	-0.093 [0.214]	-0.362 [0.285]
sex ego	-0.15 [0.212]	-0.041 [0.482]	-0.055 [0.586]	-0.23 [0.394]	0.045 [0.23]	-0.096 [0.258]	0.187 [0.381]	0.011 [0.231]	0.499 [0.212]	-0.011 [0.219]
same sex	0.487 [0.179]	0.773 [0.561]	1.56 [0.614]	-0.063 [0.276]	0.948 [0.232]	0.173 [0.17]	0.842 [0.275]	0.124 [0.265]	0.681 [0.196]	0.9 [0.332]
grade alter	-0.148 [0.171]	0.215 [0.23]	0.294 [0.26]	0.204 [0.193]	-0.006 [0.123]	-0.075 [0.1]	-0.129 [0.319]	0.457 [0.181]	0.173 [0.169]	0.184 [0.147]
grade ego	-0.363 [0.202]	-0.296 [0.243]	-0.56 [0.458]	-0.099 [0.185]	-0.084 [0.122]	-0.049 [0.069]	-0.311 [0.31]	-0.017 [0.138]	0.149 [0.16]	0.228 [0.145]
grade similarity	0.345 [0.37]	0.824 [0.797]	0.188 [0.615]	-0.203 [0.546]	0.268 [0.636]	0.882 [0.357]	0.586 [0.476]	0.856 [0.459]	-0.591 [0.328]	-1.416 [0.685]
PGG alter	-0.008 [0.095]	1.199 [0.594]	0.402 [0.285]	0.058 [0.157]	0.007 [0.101]	-0.064 [0.105]	0.156 [0.177]	-0.159 [0.13]	0.087 [0.09]	-0.039 [0.132]
PGG ego	0.253 [0.151]	0.213 [0.424]	-0.099 [0.256]	-0.362 [0.216]	-0.138 [0.094]	0.059 [0.092]	0.111 [0.196]	-0.073 [0.111]	-0.016 [0.086]	-0.312 [0.12]
PGG similarity	-0.338 [0.438]	-1.513 [0.895]	-0.074 [0.839]	0.325 [0.535]	0.333 [0.354]	-0.717 [0.351]	1.204 [0.56]	-0.28 [0.477]	0.015 [0.285]	-0.817 [0.486]
popularity alter	0.087 [0.066]	0.267 [0.221]	0.069 [0.165]	0.229 [0.126]	0.235 [0.084]	0.096 [0.068]	0.418 [0.188]	0.378 [0.134]	0.191 [0.075]	0.293 [0.109]
popularity ego	0.191 [0.086]	-0.092 [0.203]	0.309 [0.211]	-0.008 [0.1]	-0.017 [0.06]	-0.016 [0.048]	0.138 [0.166]	-0.031 [0.06]	0.099 [0.053]	0.017 [0.079]
popularity ego x popularity alter	0.019 [0.014]	0.292 [0.205]	-0.164 [0.083]	0.059 [0.038]	0.033 [0.03]	0.036 [0.018]	-0.056 [0.055]	0.026 [0.022]	-0.027 [0.02]	0.059 [0.034]
Overall maximum convergence ratio:	0.156	0.193	0.109	0.195	0.209	0.203	0.207	0.217	0.156	0.193

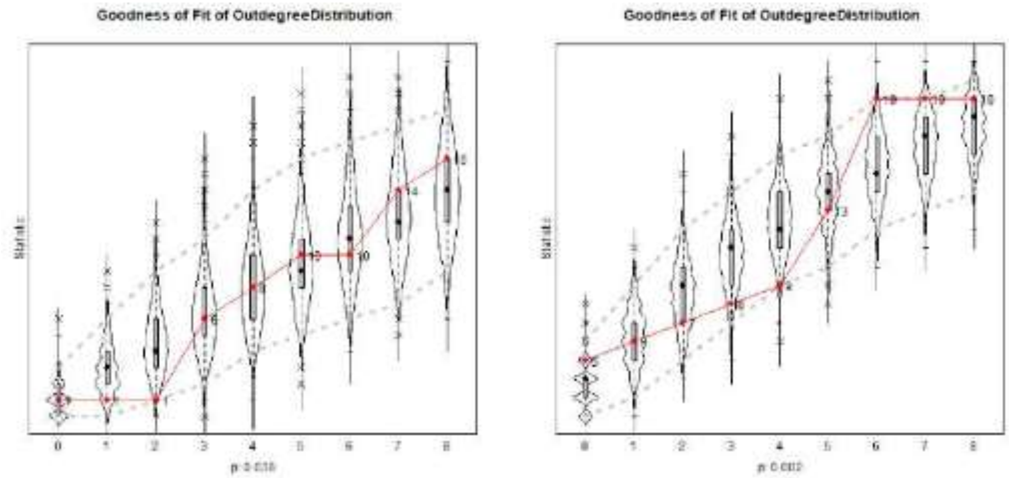
	11	12	13	14	15	16	17	18	19	20
outdegree (density)	-3.069	-2.333	-1.886	-1.233	-0.009	-0.927	1.304	-0.38	-1.754	-2.83
	[0.822]	[0.79]	[0.691]	[0.696]	[1.169]	[0.624]	[2.771]	[1.7]	[2.119]	[3.196]
reciprocity	0.625	1.575	1.189	1.259	0.758	0.51	1.285	2.011	1.583	1.866
	[0.656]	[0.687]	[0.573]	[0.436]	[0.473]	[0.391]	[0.594]	[1.187]	[0.983]	[1.442]
transitive triplets	0.086	0.101	0.364	0.272	0.206	0.08	0.397	1.315	0.224	-0.176
	[0.145]	[0.08]	[0.176]	[0.138]	[0.074]	[0.083]	[0.196]	[0.615]	[0.29]	[0.436]
transitive recipr. triplets	0.077	-0.088	-0.265	-0.271	-0.158	0.167	0.032	-0.999	-0.114	0.322
	[0.187]	[0.121]	[0.268]	[0.192]	[0.084]	[0.111]	[0.315]	[0.882]	[0.456]	[0.757]
indegree - popularity	-0.087	0.056	0.087	-0.104	-0.244	-0.018	-0.12	-0.026	-0.303	0.362
	[0.123]	[0.073]	[0.084]	[0.1]	[0.146]	[0.05]	[0.256]	[0.148]	[0.276]	[0.243]
outdegree - popularity	-0.021	-0.085	-0.105	-0.022	0.087	-0.125	0.126	0.188	-0.157	0
	[0.077]	[0.046]	[0.113]	[0.092]	[0.061]	[0.041]	[0.865]	[0.564]	[0.14]	[NA]
outdegree - activity	0.108	0.067	-0.003	-0.004	0.012	0.003	-0.526	-0.687	0.195	-0.212
	[0.037]	[0.026]	[0.045]	[0.038]	[0.025]	[0.027]	[0.228]	[0.426]	[0.124]	[0.342]
sex alter	0.549	0.168	0.871	-0.005	-0.434	0.323	0.286	0.098	-0.162	0.423
	[0.354]	[0.228]	[0.351]	[0.241]	[0.306]	[0.217]	[0.61]	[0.546]	[0.885]	[0.831]
sex ego	-0.221	0.409	0.841	0.191	-0.363	0.02	-0.855	-0.358	-1.306	-2.119
	[0.242]	[0.274]	[0.405]	[0.235]	[0.169]	[0.213]	[0.691]	[0.721]	[1.318]	[1.941]
same sex	1.328	0.379	0.248	0.889	0.227	0.427	0.565	0.852	0.318	1.968
o	[0.325]	[0.221]	[0.362]	[0.239]	[0.154]	[0.18]	[0.359]	[0.472]	[0.598]	[0.986]
grade alter	-0.149	-0.229	0.032	0.31	-0.104	0.149	0.338	0.038	0.286	0.284
	[0.178]	[0.193]	[0.12]	[0.175]	[0.109]	[0.117]	[0.213]	[0.29]	[0.404]	[0.333]
grade ego	-0.261	0.326	0.028	-0.064	-0.229	0.15	0.197	0.077	0.4	0.3
	[0.155]	[0.229]	[0.157]	[0.121]	[0.095]	[0.135]	[0.29]	[0.412]	[0.437]	[0.565]
grade similarity	1.292	-0.264	0.978	-0.159	0.362	-0.1	-0.369	1.09	-1.567	0.85
	[0.598]	[0.407]	[0.663]	[0.483]	[0.389]	[0.384]	[0.946]	[0.882]	[1.25]	[1.245]
PGG alter	0.048	-0.005	0.04	-0.005	0.36	0.046	0.146	-0.138	0.036	-0.449
	[0.096]	[0.096]	[0.147]	[0.091]	[0.16]	[0.077]	[0.213]	[0.138]	[0.242]	[0.444]
PGG ego	-0.076	0.199	-0.058	0.029	0.173	0.143	0.256	0.084	-0.678	0.697
	[0.073]	[0.138]	[0.163]	[0.09]	[0.075]	[0.07]	[0.276]	[0.174]	[0.36]	[0.86]
PGG similarity	0.411	-0.583	0.522	0.049	-0.047	-0.26	0.922	1.179	0.872	2.114
	[0.379]	[0.401]	[0.658]	[0.324]	[0.293]	[0.31]	[0.617]	[0.734]	[1.055]	[1.852]
popularity alter	0.459	0.199	-0.029	0.132	0.401	0.263	0.168	-0.013	0.402	0.218
	[0.162]	[0.082]	[0.096]	[0.091]	[0.206]	[0.092]	[0.197]	[0.185]	[0.22]	[0.289]
popularity ego	-0.012	0.085	-0.199	0.04	0.051	-0.081	-0.136	-0.285	-0.458	-0.146
	[0.091]	[0.075]	[0.13]	[0.073]	[0.063]	[0.062]	[0.198]	[0.286]	[0.249]	[0.334]
popularity ego x popularity alter	-0.019	0.024	0.08	0.027	0.057	0.025	0.075	-0.004	0.062	0.086
	[0.028]	[0.022]	[0.045]	[0.023]	[0.024]	[0.032]	[0.053]	[0.07]	[0.055]	[0.081]
Overall maximum convergence ratio:	0.219	0.179	0.173	0.182	0.144	0.146	0.176	0.115	0.244	0.116

Table 5: Results of the meta-analysis for friendship network and PGG, Model 2

	est	se	N	p	tau2	Q	Qp
outdegree (density)	-1.247	0.295	20.000	0.000	0.803	33.420	0.021
reciprocity	1.070	0.127	20.000	0.000	0.000	9.857	0.956
transitive triplets	0.178	0.029	20.000	0.000	0.000	14.764	0.737
transitive recipr. triplets	-0.045	0.043	20.000	0.295	0.045	14.077	0.779
indegree - popularity	-0.022	0.020	20.000	0.273	0.002	20.704	0.353
outdegree - popularity	-0.102	0.020	19.000	0.000	0.041	21.945	0.234
outdegree - activity	0.026	0.012	20.000	0.029	0.031	37.279	0.007
sex alter	0.012	0.099	20.000	0.906	0.282	32.188	0.030
sex ego	0.011	0.076	20.000	0.884	0.153	23.363	0.222
same sex	0.566	0.087	20.000	0.000	0.254	36.868	0.008
grade alter	0.072	0.042	20.000	0.089	0.070	21.335	0.319
grade ego	-0.036	0.042	20.000	0.396	0.083	23.531	0.215
grade similarity	0.219	0.148	20.000	0.139	0.379	29.760	0.055
PGG alter	0.018	0.036	20.000	0.623	0.002	18.877	0.465
PGG ego	0.019	0.035	20.000	0.592	0.076	27.212	0.100
PGG similarity	-0.018	0.111	20.000	0.872	0.169	25.313	0.150
popularity alter	0.184	0.029	20.000	0.000	0.051	21.689	0.300
popularity ego	0.009	0.022	20.000	0.668	0.037	24.560	0.176
popularity ego x popularity alter	0.023	0.007	20.000	0.002	0.013	27.587	0.092
int. PGG ego x PGG similarity	0.049	0.101	20.000	0.629	0.000	17.652	0.546

Table 6: Goodness of fit statistics, Friendship and PGG, Model 2

	Indegree distribution	Outdegree distribution	Geodesic Distance	Triad Census
1	0.759	0.686	0.968	0.994
2	0.910	0.077	0.557	0.977
3	0.866	0.747	0.291	0.624
4	0.748	0.402	0.847	0.994
5	0.456	0.038	0.290	0.162
6	0.660	0.759	0.349	0.246
7	0.233	0.671	0.713	0.858
8	0.386	0.768	0.987	0.979
9	0.883	0.931	0.583	0.963
10	0.848	0.553	0.489	0.989
11	0.469	0.002	0.325	0.835
12	0.216	0.666	0.762	0.822
13	0.226	0.328	0.144	0.652
14	0.846	0.810	0.947	0.893
15	0.804	0.971	0.871	0.853
16	0.878	0.468	0.397	0.241
17	0.612	0.814	0.994	0.993
18	0.991	0.514	0.926	0.965
19	0.985	0.697	0.352	0.673
20	0.743	0.649	0.946	0.766



(a) Class 5

(b) Class 11

Figure 2: Classes with poor fit, Friendship and PGG, Model 2

Table 7: Results of separate SAOMs for friendship network and PGG, Model 2

	1	2	3	4	5	6	7	8	9	10
outdegree (density)	-0.94	0.466	-5.541	1.112	-2.828	-0.252	-0.078	1.487	-1.289	-1.654
	[0.628]	[1.878]	[1.816]	[1.44]	[0.806]	[0.886]	[0.902]	[1.181]	[0.642]	[0.942]
reciprocity	1.158	1.136	1.727	0.584	1.779	0.503	0.642	1.086	1.044	1.888
	[0.385]	[1.016]	[0.873]	[0.619]	[0.576]	[0.444]	[0.696]	[0.397]	[0.41]	[0.787]
transitive triplets	0.219	0.428	0.197	0.028	-0.019	0.106	0.266	0.299	0.176	0.308
	[0.093]	[0.301]	[0.196]	[0.177]	[0.116]	[0.128]	[0.152]	[0.115]	[0.083]	[0.132]
transitive recipr. triplets	0.033	-0.2	-0.217	0.212	0.228	0.046	-0.038	-0.082	-0.085	-0.255
	[0.131]	[0.367]	[0.287]	[0.246]	[0.191]	[0.196]	[0.261]	[0.18]	[0.14]	[0.197]
indegree - popularity	0.02	-0.112	0.108	-0.09	0.058	-0.102	-0.116	-0.338	-0.06	-0.11
	[0.045]	[0.171]	[0.133]	[0.119]	[0.074]	[0.096]	[0.092]	[0.15]	[0.063]	[0.095]
outdegree - popularity	-0.198	-0.231	0.027	-0.123	-0.151	-0.091	-0.133	-0.283	-0.099	-0.133
	[0.066]	[0.233]	[0.102]	[0.16]	[0.053]	[0.056]	[0.096]	[0.083]	[0.036]	[0.052]
outdegree - activity	-0.04	-0.081	0.232	-0.086	0.094	0.023	-0.088	0.009	0.029	0.068
	[0.034]	[0.115]	[0.108]	[0.089]	[0.032]	[0.03]	[0.062]	[0.034]	[0.022]	[0.036]
sex alter	-0.035	0.139	-0.467	-0.579	0.356	0.232	-0.177	-1.068	-0.08	-0.365
	[0.184]	[0.483]	[0.532]	[0.368]	[0.23]	[0.289]	[0.376]	[0.341]	[0.22]	[0.305]
sex ego	-0.135	0.02	-0.188	-0.164	0.044	-0.092	0.171	0.024	0.478	-0.011
	[0.222]	[0.539]	[0.584]	[0.373]	[0.221]	[0.256]	[0.399]	[0.229]	[0.201]	[0.229]
same sex	0.492	0.736	1.474	-0.041	0.978	0.165	0.844	0.127	0.674	0.902
	[0.176]	[0.569]	[0.583]	[0.278]	[0.253]	[0.171]	[0.273]	[0.271]	[0.199]	[0.321]
grade alter	-0.147	0.219	0.302	0.222	-0.021	-0.077	-0.136	0.463	0.156	0.176
	[0.17]	[0.24]	[0.24]	[0.183]	[0.128]	[0.1]	[0.308]	[0.196]	[0.176]	[0.15]
grade ego	-0.36	-0.309	-0.502	-0.09	-0.074	-0.034	-0.315	-0.032	0.164	0.228
	[0.196]	[0.245]	[0.435]	[0.187]	[0.128]	[0.074]	[0.314]	[0.131]	[0.159]	[0.139]
grade similarity	0.338	0.814	0.181	-0.188	0.311	0.882	0.577	0.859	-0.578	-1.401
	[0.348]	[0.775]	[0.626]	[0.563]	[0.631]	[0.356]	[0.448]	[0.458]	[0.321]	[0.693]
PGG alter	0.002	1.325	0.747	-0.04	-0.112	-0.002	0.182	-0.279	-0.013	-0.047
	[0.113]	[0.69]	[0.413]	[0.182]	[0.136]	[0.118]	[0.182]	[0.157]	[0.146]	[0.143]
PGG ego	0.241	0.151	-0.237	-0.323	-0.089	0.023	0.039	0.033	0.005	-0.303
	[0.139]	[0.487]	[0.265]	[0.219]	[0.109]	[0.092]	[0.261]	[0.122]	[0.091]	[0.149]
PGG similarity	-0.32	-1.407	0.638	0.304	0.355	-0.694	1.288	-0.532	-0.076	-0.87
	[0.415]	[0.963]	[0.996]	[0.528]	[0.365]	[0.353]	[0.607]	[0.511]	[0.288]	[0.552]
popularity alter	0.079	0.262	0.062	0.224	0.239	0.096	0.417	0.385	0.188	0.299
	[0.063]	[0.225]	[0.156]	[0.126]	[0.091]	[0.072]	[0.167]	[0.139]	[0.072]	[0.112]
popularity ego	0.191	-0.099	0.269	-0.017	-0.02	-0.021	0.144	-0.045	0.093	0.015
	[0.081]	[0.201]	[0.211]	[0.1]	[0.068]	[0.047]	[0.137]	[0.06]	[0.053]	[0.086]
popularity ego x popularity alter	0.019	0.295	-0.153	0.062	0.034	0.036	-0.057	0.025	-0.026	0.06
	[0.014]	[0.226]	[0.072]	[0.039]	[0.029]	[0.017]	[0.051]	[0.022]	[0.02]	[0.037]
int. PGG ego x PGG similarity	-0.062	-0.345	-1.225	0.554	0.699	-0.324	-0.282	0.606	0.297	0.073
	[0.413]	[1.526]	[0.89]	[0.499]	[0.48]	[0.328]	[0.679]	[0.378]	[0.354]	[0.416]
Overall maximum convergence ratio:	0.204	0.156	0.130	0.142	0.223	0.167	0.173	0.128	0.196	0.184

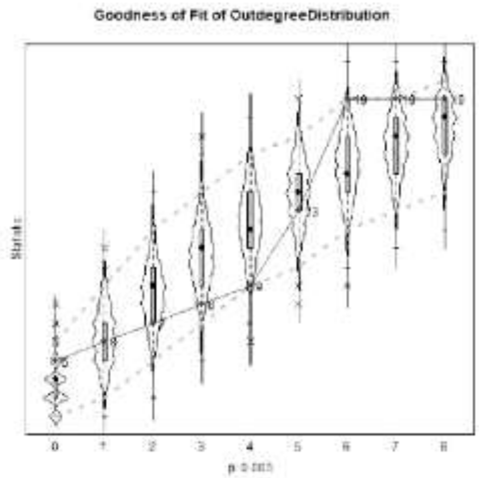
	11	12	13	14	15	16	17	18	19	20
outdegree (density)	-3.126	-2.413	-1.955	-1.403	0.082	-0.885	1.161	-0.239	-1.726	-3.018
	[0.822]	[0.76]	[0.691]	[0.705]	[1.358]	[0.617]	[2.093]	[2.014]	[2.113]	[3.548]
reciprocity	0.736	1.518	1.214	1.298	0.789	0.54	1.351	1.999	1.603	2.018
	[0.63]	[0.677]	[0.547]	[0.443]	[0.528]	[0.445]	[0.662]	[1.629]	[1.048]	[1.363]
transitive triplets	0.095	0.1	0.365	0.281	0.21	0.083	0.408	1.282	0.246	-0.251
	[0.15]	[0.095]	[0.164]	[0.134]	[0.081]	[0.084]	[0.223]	[1.018]	[0.291]	[0.441]
transitive recipr. triplets	0.06	-0.088	-0.267	-0.287	-0.163	0.162	0.015	-0.952	-0.134	0.468
	[0.191]	[0.146]	[0.257]	[0.189]	[0.091]	[0.117]	[0.354]	[1.427]	[0.439]	[0.688]
indegree - popularity	-0.069	0.061	0.088	-0.099	-0.254	-0.021	-0.101	-0.026	-0.312	0.407
	[0.12]	[0.068]	[0.082]	[0.117]	[0.161]	[0.053]	[0.235]	[0.159]	[0.279]	[0.268]
outdegree - popularity	-0.039	-0.081	-0.109	-0.02	0.089	-0.127	0.136	0.15	-0.158	0
	[0.073]	[0.045]	[0.121]	[0.09]	[0.063]	[0.04]	[0.799]	[0.931]	[0.146]	[NA]
outdegree - activity	0.109	0.067	-0.005	0.001	0.011	0.003	-0.522	-0.672	0.191	-0.254
	[0.037]	[0.025]	[0.043]	[0.037]	[0.025]	[0.026]	[0.266]	[0.568]	[0.119]	[0.391]
sex alter	0.535	0.162	0.905	-0.012	-0.447	0.328	0.33	0.133	-0.156	0.559
	[0.358]	[0.231]	[0.352]	[0.234]	[0.32]	[0.22]	[0.525]	[0.657]	[0.935]	[0.886]
sex ego	-0.231	0.417	0.822	0.202	-0.369	0.024	-0.829	-0.39	-1.328	-2.595
	[0.253]	[0.292]	[0.427]	[0.237]	[0.17]	[0.214]	[0.658]	[0.748]	[1.244]	[2.021]
same sex	1.311	0.394	0.33	0.911	0.231	0.422	0.555	0.807	0.299	2.202
	[0.307]	[0.229]	[0.36]	[0.238]	[0.157]	[0.184]	[0.347]	[0.4]	[0.635]	[1.025]
grade alter	-0.141	-0.217	0.047	0.317	-0.088	0.154	0.351	0.024	0.274	0.372
	[0.182]	[0.197]	[0.123]	[0.191]	[0.115]	[0.12]	[0.206]	[0.372]	[0.449]	[0.383]
grade ego	-0.261	0.302	0.049	-0.042	-0.222	0.162	0.17	0.066	0.433	0.311
	[0.148]	[0.229]	[0.16]	[0.124]	[0.099]	[0.139]	[0.286]	[0.467]	[0.467]	[0.576]
grade similarity	1.255	-0.261	1.007	-0.095	0.382	-0.104	-0.318	1.035	-1.561	0.97
	[0.563]	[0.415]	[0.681]	[0.477]	[0.402]	[0.396]	[0.919]	[0.984]	[1.201]	[1.384]
PGG alter	0.087	0.072	0.232	-0.111	0.296	0.087	-0.016	-0.006	0.09	-0.702
	[0.141]	[0.136]	[0.193]	[0.135]	[0.213]	[0.1]	[0.248]	[0.37]	[0.278]	[0.553]
PGG ego	-0.079	0.176	-0.135	0.047	0.181	0.128	0.22	0.084	-0.737	0.888
	[0.076]	[0.153]	[0.162]	[0.093]	[0.076]	[0.077]	[0.286]	[0.194]	[0.393]	[0.869]
PGG similarity	0.454	-0.429	0.375	0.131	-0.093	-0.258	0.746	1.187	0.719	2.341
	[0.377]	[0.447]	[0.658]	[0.334]	[0.312]	[0.31]	[0.646]	[0.789]	[1.308]	[1.952]
popularity alter	0.45	0.195	-0.035	0.13	0.409	0.267	0.154	-0.009	0.409	0.238
	[0.171]	[0.083]	[0.1]	[0.1]	[0.225]	[0.099]	[0.218]	[0.203]	[0.226]	[0.32]
popularity ego	-0.016	0.096	-0.211	0.039	0.051	-0.086	-0.13	-0.284	-0.448	-0.166
	[0.098]	[0.087]	[0.127]	[0.072]	[0.063]	[0.064]	[0.181]	[0.268]	[0.236]	[0.353]
popularity ego x popularity alter	-0.019	0.022	0.09	0.024	0.057	0.027	0.084	0	0.062	0.09
	[0.028]	[0.021]	[0.047]	[0.024]	[0.025]	[0.033]	[0.054]	[0.078]	[0.054]	[0.089]
int. PGG ego x PGG similarity	-0.154	-0.284	-1.285	0.373	0.26	-0.165	0.649	-0.326	-0.41	0.864
	[0.369]	[0.383]	[0.658]	[0.351]	[0.363]	[0.293]	[0.65]	[0.745]	[1.349]	[1.51]
Overall maximum convergence ratio:	0.204	0.179	0.124	0.167	0.193	0.179	0.167	0.197	0.136	0.173

Table 8: Results of the meta-analysis for friendship network and DG, Model 1

	est	se	N	p	tau2	Q	Qp
outdegree (density)	-1.264	0.2926	20	0.000	0.773	32.013	0.031
reciprocity	1.081	0.1254	20	0.000	0.000	10.242	0.947
transitive triplets	0.176	0.0282	20	0.000	0.002	14.303	0.766
transitive recipr. triplets	-0.051	0.0459	20	0.271	0.075	17.035	0.587
indegree - popularity	-0.015	0.0212	20	0.489	0.000	20.871	0.344
outdegree - popularity	-0.110	0.0156	19	0.000	0.000	14.854	0.672
outdegree - activity	0.027	0.0126	20	0.034	0.034	36.108	0.010
sex alter	0.054	0.0949	20	0.571	0.249	30.492	0.046
sex ego	0.038	0.0719	20	0.596	0.142	22.934	0.240
same sex	0.554	0.0867	20	0.000	0.257	37.208	0.007
grade alter	0.058	0.0374	20	0.119	0.000	15.528	0.688
grade ego	-0.041	0.0398	20	0.304	0.074	22.730	0.249
grade similarity	0.232	0.1511	20	0.125	0.412	32.290	0.029
DG alter	-0.003	0.0190	20	0.885	0.000	18.290	0.503
DG ego	-0.013	0.0193	20	0.504	0.034	20.590	0.360
DG similarity	-0.139	0.0852	20	0.102	0.000	12.999	0.839
popularity alter	0.195	0.0286	20	0.000	0.046	19.372	0.433
popularity ego	-0.004	0.0190	20	0.844	0.027	21.314	0.320
popularity ego x popularity alter	0.024	0.0068	20	0.000	0.010	24.887	0.164

Table 9: Goodness of fit statistics, Friendship and DG, Model 1

	Indegree distribution	Outdegree distribution	Geodesic Distance	Triad Census
1	0.832	0.643	0.954	0.986
2	0.866	0.464	0.440	0.861
3	0.836	0.642	0.213	0.575
4	0.754	0.352	0.896	0.994
5	0.494	0.074	0.349	0.227
6	0.719	0.776	0.392	0.205
7	0.051	0.563	0.630	0.648
8	0.404	0.649	0.984	0.958
9	0.842	0.951	0.579	0.944
10	0.824	0.716	0.521	0.977
11	0.524	0.003	0.347	0.836
12	0.192	0.461	0.713	0.783
13	0.229	0.272	0.168	0.767
14	0.887	0.762	0.936	0.866
15	0.942	0.962	0.950	0.870
16	0.851	0.522	0.357	0.258
17	0.592	0.754	0.996	0.996
18	0.996	0.580	0.890	0.955
19	0.989	0.658	0.451	0.741
20	0.791	0.745	0.947	0.742



(a) Class 11

Figure 3: Classes with poor fit, Friendship and DG, Model 1

Table 10: Results of separate SAOMs for friendship network and DG, Model 1

		1	2	3	4	5	6	7	8	9	10
	outdegree (density)	-0.893	-0.48	-5.373	0.92	-2.771	-0.214	0.3	2.42	-1.072	-1.789
		[0.573]	[1.849]	[2.063]	[1.359]	[0.78]	[0.922]	[1.097]	[1.388]	[0.633]	[0.96]
	reciprocity	1.092	0.966	1.738	0.676	1.7	0.594	0.675	1.122	1.018	1.752
		[0.375]	[0.849]	[0.966]	[0.643]	[0.585]	[0.423]	[0.688]	[0.412]	[0.374]	[0.663]
	transitive triplets	0.21	0.218	0.176	0.059	-0.016	0.109	0.256	0.313	0.188	0.269
		[0.088]	[0.265]	[0.221]	[0.167]	[0.105]	[0.123]	[0.145]	[0.119]	[0.081]	[0.132]
	transitive recipr. triplets	0.051	-0.103	-0.261	0.168	0.232	0.029	0.005	-0.114	-0.072	-0.231
		[0.135]	[0.31]	[0.311]	[0.222]	[0.163]	[0.191]	[0.243]	[0.193]	[0.141]	[0.188]
	indegree - popularity	0.012	0.112	0.125	-0.133	0.052	-0.108	-0.263	-0.443	-0.087	-0.091
		[0.043]	[0.148]	[0.132]	[0.134]	[0.082]	[0.121]	[0.207]	[0.178]	[0.07]	[0.096]
	outdegree - popularity	-0.188	-0.161	0.03	-0.095	-0.144	-0.076	-0.044	-0.348	-0.107	-0.12
		[0.06]	[0.251]	[0.1]	[0.159]	[0.053]	[0.064]	[0.13]	[0.102]	[0.036]	[0.043]
	outdegree - activity	-0.038	-0.073	0.249	-0.062	0.093	0.015	-0.084	0.006	0.024	0.07
		[0.033]	[0.124]	[0.141]	[0.077]	[0.029]	[0.031]	[0.057]	[0.034]	[0.023]	[0.032]
	sex alter	-0.042	0.496	0.215	-0.559	0.461	0.251	-0.59	-1.274	-0.079	-0.295
		[0.177]	[0.417]	[0.537]	[0.381]	[0.228]	[0.307]	[0.598]	[0.417]	[0.217]	[0.233]
	sex ego	-0.121	0.069	-0.419	-0.313	0.046	-0.082	0.13	0.068	0.514	0.135
		[0.214]	[0.458]	[0.638]	[0.369]	[0.234]	[0.199]	[0.355]	[0.231]	[0.212]	[0.183]
cc	same sex	0.48	0.633	1.399	-0.069	0.954	0.107	0.83	0.1	0.638	0.924
		[0.17]	[0.495]	[0.608]	[0.275]	[0.236]	[0.159]	[0.271]	[0.271]	[0.199]	[0.282]
	grade alter	-0.111		0.026	0.192	0.094	-0.071	0.011	0.464	0.188	0.164
		[0.178]	0 [0.185]	[0.204]	[0.164]	[0.139]	[0.106]	[0.362]	[0.214]	[0.17]	[0.147]
	grade ego	-0.35	-0.35	-0.411	0.103	-0.199	-0.048	-0.372	-0.015	0.134	0.155
		[0.199]	[0.269]	[0.474]	[0.178]	[0.127]	[0.066]	[0.274]	[0.147]	[0.158]	[0.11]
	grade similarity	0.37	0.586	0.117	-0.154	0.978	0.797	0.681	0.868	-0.659	-1.186
		[0.334]	[0.655]	[0.651]	[0.542]	[0.684]	[0.333]	[0.437]	[0.459]	[0.321]	[0.598]
	DG alter	0.053	0.098	-0.064	0.165	-0.161	-0.046	0.431	-0.164	-0.077	-0.05
		[0.054]	[0.119]	[0.118]	[0.153]	[0.112]	[0.098]	[0.301]	[0.1]	[0.062]	[0.074]
	DG ego	-0.036	0.25	0.116	-0.148	-0.004	0.068	-0.16	0.011	-0.096	-0.061
		[0.059]	[0.167]	[0.152]	[0.13]	[0.104]	[0.049]	[0.132]	[0.071]	[0.056]	[0.054]
	DG similarity	-0.395	0.004	0.78	-0.502	-0.236	0.086	0.415	-0.317	-0.215	-0.691
		[0.291]	[0.547]	[0.696]	[0.544]	[0.499]	[0.23]	[0.542]	[0.363]	[0.316]	[0.409]
	popularity alter	0.081	0.08	0.177	0.257	0.252	0.124	0.359	0.463	0.235	0.279
		[0.059]	[0.192]	[0.15]	[0.131]	[0.092]	[0.099]	[0.19]	[0.165]	[0.09]	[0.103]
	popularity ego	0.108	-0.236	0.264	-0.042	0.023	-0.029	0.208	-0.048	0.115	0.012
		[0.069]	[0.217]	[0.243]	[0.098]	[0.066]	[0.052]	[0.161]	[0.059]	[0.055]	[0.069]
	popularity ego x popularity alter	0.015	0.226	-0.16	0.057	0.035	0.034	-0.053	0.027	-0.025	0.055
		[0.013]	[0.171]	[0.084]	[0.038]	[0.034]	[0.017]	[0.06]	[0.022]	[0.021]	[0.032]
	Overall maximum convergence ratio:	0.152	0.107	0.105	0.178	0.191	0.177	0.185	0.156	0.143	0.176

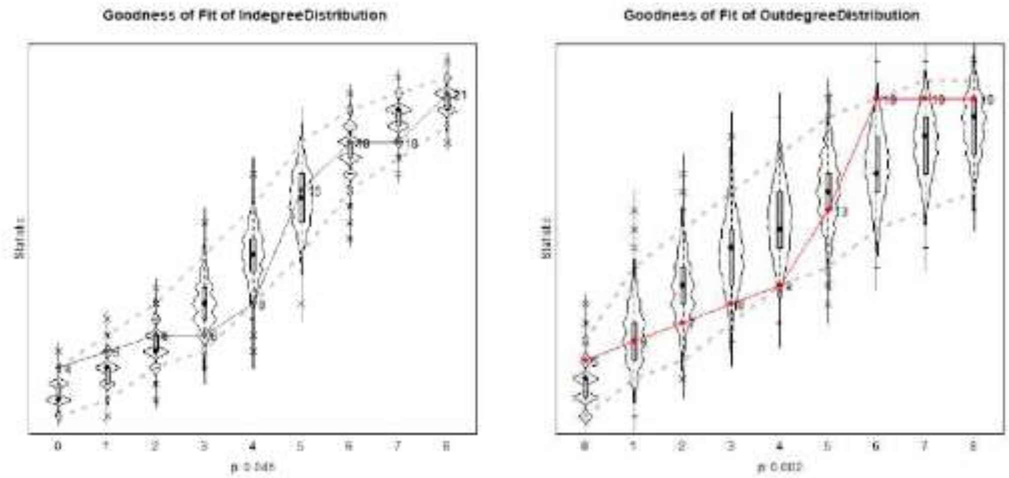
	11	12	13	14	15	16	17	18	19	20
outdegree (density)	-2.989	-2.845	-1.804	-1.264	-0.442	-1.03	0.729	-0.289	-1.388	-2.619
	[0.804]	[0.811]	[0.685]	[0.734]	[1.567]	[0.615]	[3.373]	[1.459]	[1.762]	[2.991]
reciprocity	0.766	1.932	1.146	1.305	0.963	0.495	1.501	2.08	1.619	1.742
	[0.718]	[0.708]	[0.574]	[0.412]	[0.564]	[0.406]	[0.844]	[1.252]	[1.026]	[1.283]
transitive triplets	0.082	0.139	0.345	0.281	0.216	0.07	0.408	1.335	0.355	-0.07
	[0.14]	[0.093]	[0.188]	[0.121]	[0.086]	[0.085]	[0.269]	[0.734]	[0.287]	[0.493]
transitive recipr. triplets	0.06	-0.168	-0.239	-0.292	-0.19	0.181	0.132	-0.983	-0.425	-0.063
	[0.184]	[0.135]	[0.273]	[0.179]	[0.096]	[0.123]	[0.384]	[0.964]	[0.512]	[0.54]
indegree - popularity	-0.066	0.077	0.068	-0.094	-0.239	-0.016	-0.433	-0.048	-0.301	0.353
	[0.126]	[0.066]	[0.091]	[0.104]	[0.291]	[0.054]	[0.832]	[0.143]	[0.274]	[0.248]
outdegree - popularity	-0.041	-0.085	-0.1	-0.029	0.131	-0.121	1.196	0.26	-0.112	0
	[0.085]	[0.042]	[0.12]	[0.091]	[0.16]	[0.039]	[2.314]	[0.525]	[0.133]	[NA]
outdegree - activity	0.106	0.081	-0.002	-0.006	0.008	0.011	-0.837	-0.725	0.13	-0.149
	[0.037]	[0.029]	[0.046]	[0.036]	[0.025]	[0.026]	[0.648]	[0.5]	[0.076]	[0.326]
sex alter	0.495	0.142	0.765	0.009	-0.304	0.295	1.321	0.083	0.198	0.115
	[0.332]	[0.225]	[0.342]	[0.235]	[0.388]	[0.203]	[2.217]	[0.55]	[0.88]	[0.768]
sex ego	-0.237	0.373	0.733	0.256	-0.366	0.088	-1.129	-0.575	-0.584	-1.413
	[0.245]	[0.267]	[0.374]	[0.245]	[0.191]	[0.202]	[1.395]	[1.003]	[0.684]	[1.116]
same sex	1.265	0.468	0.309	0.921	0.238	0.401	0.577	0.815	0.74	1.966
	[0.321]	[0.218]	[0.359]	[0.253]	[0.154]	[0.171]	[0.408]	[0.457]	[0.604]	[1.074]
grade alter	-0.148	-0.195	0.018	0.3	-0.056	0.16	0.607	0.064	0.192	0.158
	[0.184]	[0.2]	[0.127]	[0.169]	[0.11]	[0.115]	[0.849]	[0.324]	[0.363]	[0.297]
grade ego	-0.24	0.392	0.014	-0.035	-0.172	0.118	-0.109	-0.116	0.195	0.204
	[0.145]	[0.23]	[0.149]	[0.136]	[0.09]	[0.128]	[0.447]	[0.431]	[0.279]	[0.461]
grade similarity	1.197	-0.206	1.225	-0.091	0.116	-0.098	-0.2	0.928	-1.534	0.711
	[0.585]	[0.405]	[0.67]	[0.517]	[0.408]	[0.378]	[1.086]	[0.777]	[1.128]	[1.005]
DG alter	0.044	0.035	-0.136	-0.006	0.259	0.026	0.614	0.027	0.065	0.12
	[0.079]	[0.06]	[0.081]	[0.069]	[0.24]	[0.043]	[0.923]	[0.127]	[0.13]	[0.181]
DG ego	0.023	-0.118	-0.051	-0.016	-0.042	0.075	-0.54	-0.057	0.06	-0.029
	[0.061]	[0.073]	[0.088]	[0.067]	[0.065]	[0.043]	[0.398]	[0.156]	[0.112]	[0.23]
DG similarity	0.142	-0.07	0.196	-0.621	-0.268	0.001	0.44	-0.166	-0.826	-0.606
	[0.391]	[0.32]	[0.389]	[0.356]	[0.344]	[0.232]	[0.645]	[0.554]	[0.684]	[0.789]
popularity alter	0.454	0.203	-0.003	0.125	0.407	0.259	0.254	0.049	0.48	0.306
	[0.164]	[0.075]	[0.101]	[0.088]	[0.432]	[0.095]	[0.305]	[0.198]	[0.228]	[0.257]
popularity ego	-0.019	-0.026	-0.181	0.038	-0.001	-0.108	0.009	-0.166	-0.223	-0.155
	[0.094]	[0.052]	[0.127]	[0.061]	[0.054]	[0.066]	[0.246]	[0.243]	[0.128]	[0.264]
popularity ego x popularity alter	-0.022	0.024	0.073	0.031	0.06	0.027	0.059	0.021	0.062	0.12
	[0.028]	[0.018]	[0.043]	[0.025]	[0.024]	[0.033]	[0.06]	[0.069]	[0.044]	[0.077]
Overall maximum convergence ratio:	0.139	0.152	0.119	0.152	0.249	0.183	0.213	0.177	0.155	0.162

Table 11: Results of the meta-analysis for friendship network and DG, Model 2

	est	se	N	p	tau2	Q	Qp
outdegree (density)	-1.269	0.2814	20	0.000	0.708	31.198	0.038
reciprocity	1.077	0.1250	20	0.000	0.000	10.414	0.942
transitive triplets	0.180	0.0283	20	0.000	0.002	14.182	0.773
transitive recipr. triplets	-0.057	0.0456	20	0.214	0.069	16.799	0.603
indegree - popularity	-0.017	0.0210	20	0.410	0.000	21.775	0.296
outdegree - popularity	-0.106	0.0160	19	0.000	0.000	17.881	0.463
outdegree - activity	0.028	0.0121	20	0.022	0.031	32.566	0.027
sex alter	0.059	0.0899	20	0.512	0.217	28.531	0.074
sex ego	0.052	0.0733	20	0.479	0.142	21.730	0.298
same sex	0.557	0.0909	20	0.000	0.269	37.818	0.006
grade alter	0.065	0.0375	20	0.084	0.018	15.741	0.675
grade ego	-0.034	0.0378	20	0.367	0.055	22.626	0.254
grade similarity	0.207	0.1480	20	0.163	0.375	29.231	0.062
DG alter	0.034	0.0313	20	0.284	0.043	17.916	0.528
DG ego	-0.017	0.0195	20	0.373	0.029	18.172	0.511
DG similarity	-0.139	0.0914	20	0.127	0.059	18.549	0.486
popularity alter	0.200	0.0283	20	0.000	0.041	19.202	0.444
popularity ego	-0.003	0.0197	20	0.863	0.029	20.848	0.345
popularity ego x popularity alter	0.025	0.0071	20	0.001	0.012	25.850	0.134
int. DG ego x DG similarity	-0.129	0.0782	20	0.098	0.111	22.698	0.251

Table 12: Goodness of fit statistics, Friendship and DG, Model 2

	Indegree distribution	Outdegree distribution	Geodesic Distance	Triad Census
1	0.813	0.587	0.951	0.992
2	0.875	0.591	0.449	0.846
3	0.802	0.687	0.237	0.642
4	0.711	0.269	0.853	0.996
5	0.610	0.079	0.385	0.241
6	0.690	0.789	0.354	0.232
7	0.046	0.534	0.684	0.604
8	0.469	0.656	0.974	0.938
9	0.854	0.962	0.605	0.938
10	0.835	0.651	0.550	0.985
11	0.511	0.002	0.372	0.792
12	0.196	0.412	0.721	0.767
13	0.259	0.293	0.205	0.835
14	0.897	0.764	0.932	0.857
15	0.931	0.973	0.940	0.809
16	0.884	0.528	0.376	0.291
17	0.568	0.756	0.993	0.994
18	0.992	0.480	0.868	0.964
19	0.989	0.708	0.462	0.720
20	0.767	0.724	0.956	0.660



(a) Class 7

(b) Class 11

Figure 4: Classes with poor fit, Friendship and DG, Model 2

Table 13: Results of separate SAOMs for friendship network and DG, Model 2

	1	2	3	4	5	6	7	8	9	10
outdegree (density)	-0.852	-0.281	-5.371	0.936	-2.33	-0.22	0.091	2.486	-1.068	-1.675
	[0.586]	[2.033]	[1.949]	[1.359]	[0.781]	[0.833]	[1.091]	[1.482]	[0.686]	[0.864]
reciprocity	1.09	1.002	1.762	0.637	1.661	0.628	0.643	1.132	1.031	1.695
	[0.386]	[0.826]	[0.93]	[0.602]	[0.627]	[0.436]	[0.662]	[0.411]	[0.383]	[0.64]
transitive triplets	0.209	0.222	0.193	0.039	-0.001	0.125	0.238	0.322	0.185	0.277
	[0.095]	[0.241]	[0.213]	[0.164]	[0.119]	[0.123]	[0.161]	[0.125]	[0.087]	[0.119]
transitive recipr. triplets	0.06	-0.101	-0.283	0.176	0.253	0.015	0.029	-0.129	-0.069	-0.238
	[0.138]	[0.308]	[0.294]	[0.203]	[0.204]	[0.198]	[0.264]	[0.193]	[0.145]	[0.179]
indegree - popularity	0.01	0.109	0.122	-0.147	0.061	-0.114	-0.249	-0.448	-0.087	-0.095
	[0.044]	[0.139]	[0.146]	[0.141]	[0.082]	[0.105]	[0.192]	[0.187]	[0.069]	[0.088]
outdegree - popularity	-0.195	-0.188	0.03	-0.068	-0.16	-0.07	-0.037	-0.348	-0.107	-0.118
	[0.066]	[0.265]	[0.101]	[0.148]	[0.061]	[0.062]	[0.162]	[0.104]	[0.038]	[0.045]
outdegree - activity	-0.038	-0.078	0.249	-0.065	0.082	0.012	-0.078	0.003	0.024	0.064
	[0.034]	[0.122]	[0.127]	[0.082]	[0.032]	[0.031]	[0.067]	[0.036]	[0.024]	[0.03]
sex alter	-0.035	0.486	0.181	-0.564	0.399	0.279	-0.551	-1.274	-0.077	-0.281
	[0.174]	[0.425]	[0.47]	[0.388]	[0.211]	[0.33]	[0.633]	[0.458]	[0.216]	[0.245]
sex ego	-0.13	0.077	-0.444	-0.342	0.17	-0.103	0.12	0.026	0.516	0.147
	[0.21]	[0.464]	[0.595]	[0.347]	[0.242]	[0.223]	[0.348]	[0.245]	[0.212]	[0.195]
same sex	0.483	0.604	1.39	-0.077	0.951	0.091	0.854	0.115	0.622	0.952
	[0.177]	[0.515]	[0.615]	[0.267]	[0.274]	[0.161]	[0.27]	[0.278]	[0.201]	[0.311]
grade alter	-0.113	0.003	0.008	0.206	0.15	-0.075	0.042	0.464	0.194	0.151
	[0.189]	[0.183]	[0.217]	[0.171]	[0.133]	[0.1]	[0.408]	[0.218]	[0.175]	[0.145]
grade ego	-0.359	-0.343	-0.395	0.126	-0.282	-0.062	-0.338	-0.011	0.128	0.14
	[0.205]	[0.26]	[0.392]	[0.175]	[0.165]	[0.068]	[0.306]	[0.137]	[0.151]	[0.114]
grade similarity	0.373	0.591	0.278	-0.159	0.375	0.79	0.692	0.852	-0.655	-1.2
	[0.343]	[0.684]	[0.69]	[0.532]	[0.701]	[0.327]	[0.47]	[0.466]	[0.319]	[0.633]
DG alter	0.074	0.126	0.063	0.136	0.137	0.065	0.508	-0.088	-0.11	-0.153
	[0.083]	[0.187]	[0.162]	[0.165]	[0.188]	[0.117]	[0.317]	[0.128]	[0.081]	[0.094]
DG ego	-0.042	0.246	0.102	-0.137	-0.187	0.061	-0.189	-0.009	-0.083	-0.036
	[0.062]	[0.166]	[0.149]	[0.129]	[0.15]	[0.054]	[0.152]	[0.075]	[0.062]	[0.061]
DG similarity	-0.404	-0.002	0.695	-0.458	-1.062	0.249	0.392	-0.384	-0.182	-0.767
	[0.297]	[0.557]	[0.646]	[0.548]	[0.667]	[0.26]	[0.516]	[0.395]	[0.313]	[0.39]
popularity alter	0.086	0.083	0.195	0.266	0.247	0.131	0.357	0.469	0.237	0.291
	[0.064]	[0.19]	[0.158]	[0.131]	[0.076]	[0.097]	[0.179]	[0.17]	[0.09]	[0.102]
popularity ego	0.109	-0.232	0.249	-0.043	0.035	-0.031	0.209	-0.044	0.114	0.02
	[0.071]	[0.219]	[0.227]	[0.098]	[0.068]	[0.056]	[0.172]	[0.061]	[0.055]	[0.07]
popularity ego x popularity alter	0.015	0.227	-0.164	0.056	0.049	0.034	-0.057	0.029	-0.026	0.059
	[0.013]	[0.171]	[0.085]	[0.038]	[0.035]	[0.016]	[0.059]	[0.022]	[0.02]	[0.033]
int. DG ego x DG similarity	-0.077	-0.074	-0.489	0.187	-0.983	-0.372	-0.432	-0.265	0.128	0.418
	[0.242]	[0.428]	[0.514]	[0.392]	[0.529]	[0.231]	[0.439]	[0.284]	[0.221]	[0.303]
Overall maximum convergence ratio:	0.158	0.171	0.147	0.193	0.131	0.163	0.148	0.129	0.173	0.233

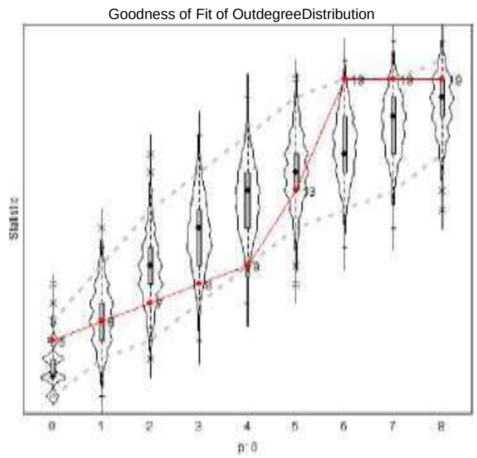
	11	12	13	14	15	16	17	18	19	20
outdegree (density)	-3.326	-2.808	-1.789	-1.294	-0.241	-1.096	1.062	-0.113	-1.353	-2.807
	[0.878]	[0.769]	[0.666]	[0.723]	[1.372]	[0.603]	[4.859]	[1.644]	[1.784]	[3.029]
reciprocity	0.774	1.943	1.172	1.302	0.943	0.534	1.481	2.233	1.684	1.824
	[0.61]	[0.623]	[0.603]	[0.449]	[0.479]	[0.398]	[1.084]	[1.287]	[0.998]	[1.381]
transitive triplets	0.07	0.132	0.363	0.277	0.22	0.065	0.4	1.379	0.384	-0.046
	[0.139]	[0.083]	[0.171]	[0.131]	[0.077]	[0.09]	[0.268]	[0.734]	[0.29]	[0.389]
transitive recipr. triplets	0.061	-0.153	-0.258	-0.27	-0.187	0.184	0.166	-1.016	-0.498	0.019
	[0.176]	[0.129]	[0.281]	[0.201]	[0.087]	[0.127]	[0.592]	[0.928]	[0.509]	[0.572]
indegree - popularity	-0.058	0.073	0.066	-0.089	-0.278	-0.015	-0.429	-0.05	-0.316	0.333
	[0.131]	[0.066]	[0.09]	[0.108]	[0.211]	[0.051]	[1.231]	[0.185]	[0.289]	[0.223]
outdegree - popularity	-0.036	-0.09	-0.101	-0.034	0.146	-0.122	1.178	0.261	-0.108	0
	[0.079]	[0.045]	[0.137]	[0.086]	[0.1]	[0.039]	[4.267]	[0.972]	[0.128]	[NA]
outdegree - activity	0.118	0.082	-0.004	-0.002	0.004	0.014	-0.869	-0.76	0.131	-0.186
	[0.038]	[0.028]	[0.042]	[0.036]	[0.026]	[0.026]	[1.179]	[0.617]	[0.077]	[0.299]
sex alter	0.541	0.152	0.747	0.01	-0.368	0.299	1.348	0.13	0.215	0.306
	[0.334]	[0.243]	[0.341]	[0.237]	[0.309]	[0.218]	[2.8]	[0.598]	[0.922]	[0.793]
sex ego	-0.142	0.356	0.726	0.247	-0.381	0.088	-1.142	-0.431	-0.622	-1.287
	[0.258]	[0.246]	[0.377]	[0.252]	[0.204]	[0.212]	[1.889]	[1.01]	[0.749]	[1.164]
same sex	1.29	0.45	0.293	0.899	0.238	0.413	0.565	0.88	0.775	2.032
	[0.342]	[0.226]	[0.381]	[0.252]	[0.15]	[0.179]	[0.766]	[0.465]	[0.593]	[0.821]
grade alter	-0.114	-0.203	0.022	0.289	-0.049	0.161	0.605	0.047	0.185	0.217
	[0.18]	[0.191]	[0.128]	[0.161]	[0.115]	[0.1]	[1.033]	[0.317]	[0.377]	[0.306]
grade ego	-0.201	0.393	0.022	-0.04	-0.153	0.12	-0.1	-0.178	0.233	0.42
	[0.149]	[0.218]	[0.157]	[0.131]	[0.096]	[0.126]	[0.379]	[0.447]	[0.298]	[0.47]
grade similarity	1.044	-0.195	1.217	-0.07	0.128	-0.076	-0.219	1.037	-1.637	0.954
	[0.613]	[0.407]	[0.701]	[0.528]	[0.409]	[0.384]	[1.18]	[0.876]	[1.145]	[1.212]
DG alter	0.218	0.064	-0.016	0.091	0.181	-0.001	0.695	0.378	0.14	-0.188
	[0.117]	[0.105]	[0.138]	[0.114]	[0.213]	[0.085]	[1.669]	[0.256]	[0.211]	[0.273]
DG ego	-0.019	-0.118	-0.069	-0.023	-0.012	0.076	-0.572	-0.102	0.061	0.085
	[0.067]	[0.065]	[0.088]	[0.067]	[0.07]	[0.046]	[1.014]	[0.157]	[0.112]	[0.263]
DG similarity	0.453	-0.044	0.2	-0.659	-0.026	0.003	0.3	-0.52	-0.915	-0.448
	[0.436]	[0.336]	[0.403]	[0.355]	[0.375]	[0.25]	[0.841]	[0.632]	[0.625]	[0.852]
popularity alter	0.453	0.208	-0.004	0.118	0.465	0.263	0.26	0.057	0.501	0.354
	[0.171]	[0.084]	[0.103]	[0.086]	[0.331]	[0.098]	[0.716]	[0.207]	[0.248]	[0.29]
popularity ego	-0.003	-0.03	-0.183	0.026		-0.115	0.008	-0.171	-0.233	-0.201
	[0.092]	[0.054]	[0.135]	[0.068]	0 [0.057]	[0.065]	[0.241]	[0.28]	[0.139]	[0.291]
popularity ego x popularity alter	-0.021	0.024	0.075	0.029	0.06	0.029	0.057	0.017	0.064	0.116
	[0.028]	[0.018]	[0.045]	[0.024]	[0.026]	[0.033]	[0.084]	[0.067]	[0.048]	[0.074]
int. DG ego x DG similarity	-0.614	-0.104	-0.415	-0.332	0.367	0.082	-0.385	-0.954	-0.213	1.354
	[0.307]	[0.28]	[0.412]	[0.311]	[0.292]	[0.225]	[0.439]	[0.562]	[0.484]	[0.888]
Overall maximum convergence ratio:	0.164	0.132	0.143	0.137	0.197	0.121	0.198	0.179	0.162	0.104

Table 14: Results of the meta-analysis for friendship network and TG, Model 1

	est	se	N	p	tau2	Q	Qp
outdegree (density)	-1.226	0.2878	20	0.000	0.759	33.638	0.020
reciprocity	1.097	0.1251	20	0.000	0.000	9.833	0.957
transitive triplets	0.175	0.0281	20	0.000	0.001	17.132	0.581
transitive recipr. triplets	-0.057	0.0473	20	0.233	0.080	16.691	0.611
indegree - popularity	-0.023	0.0211	20	0.272	0.013	24.255	0.187
outdegree - popularity	-0.101	0.0187	19	0.000	0.033	25.695	0.107
outdegree - activity	0.032	0.0112	20	0.005	0.026	37.437	0.007
sex alter	0.026	0.0915	20	0.773	0.227	32.125	0.030
sex ego	0.007	0.0668	20	0.920	0.106	20.148	0.386
same sex	0.559	0.0859	20	0.000	0.254	37.252	0.007
grade alter	0.040	0.0390	20	0.307	0.038	18.812	0.469
grade ego	-0.037	0.0402	20	0.357	0.082	25.539	0.144
grade similarity	0.199	0.1474	20	0.177	0.392	30.889	0.042
TG alter	-0.024	0.0242	20	0.311	0.063	34.548	0.016
TG ego	-0.030	0.0153	20	0.052	0.000	13.074	0.835
TG similarity	-0.081	0.1208	20	0.505	0.193	18.799	0.470
popularity alter	0.208	0.0264	20	0.000	0.021	20.013	0.394
popularity ego	-0.007	0.0178	20	0.688	0.000	16.220	0.643
popularity ego x popularity alter	0.023	0.0064	20	0.000	0.006	24.111	0.192

Table 15: Goodness of fit statistics, Friendship and TG, Model 1

	Indegree distribution	Outdegree distribution	Geodesic Distance	Triad Census
1	0.836	0.823	0.955	0.995
2	0.868	0.417	0.476	0.880
3	0.800	0.645	0.205	0.644
4	0.635	0.523	0.879	0.991
5	0.568	0.064	0.298	0.156
6	0.674	0.822	0.341	0.207
7	0.142	0.746	0.715	0.856
8	0.360	0.793	0.988	0.973
9	0.843	0.953	0.582	0.948
10	0.848	0.781	0.468	0.992
11	0.491	0.000	0.339	0.802
12	0.185	0.623	0.704	0.823
13	0.268	0.326	0.196	0.732
14	0.868	0.769	0.936	0.944
15	0.925	0.975	0.965	0.861
16	0.843	0.521	0.349	0.227
17	0.611	0.781	0.993	0.988
18	0.993	0.486	0.863	0.931
19	0.993	0.626	0.551	0.701
20	0.689	0.690	0.949	0.695



(a) Class 11

Figure 5: Classes with poor fit, Friendship and TG, Model 1

Table 16: Results of separate SAOMs for friendship network and TG, Model 1

	1	2	3	4	5	6	7	8	9	10
outdegree (density)	-0.837	-1.236	-4.849	1.059	-2.719	0.388	0.11	3.853	-1.144	-1.604
	[0.682]	[1.662]	[1.646]	[1.348]	[0.763]	[1.133]	[1.019]	[1.719]	[0.633]	[0.978]
reciprocity	1.133	0.982	1.589	0.541	1.646	0.579	0.768	1.166	1.051	1.69
	[0.385]	[0.81]	[0.783]	[0.558]	[0.576]	[0.552]	[0.845]	[0.444]	[0.384]	[0.573]
transitive triplets	0.225	0.221	0.122	0.045	-0.019	0.122	0.276	0.306	0.186	0.255
	[0.105]	[0.247]	[0.205]	[0.165]	[0.098]	[0.127]	[0.176]	[0.131]	[0.084]	[0.123]
transitive recipr. triplets	0.048	-0.107	-0.178	0.206	0.242	-0.002	-0.028	-0.089	-0.092	-0.197
	[0.146]	[0.304]	[0.29]	[0.211]	[0.185]	[0.237]	[0.31]	[0.21]	[0.155]	[0.177]
indegree - popularity	0.012	0.111	0.04	-0.121	0.061	-0.175	-0.238	-0.566	-0.088	-0.097
	[0.044]	[0.133]	[0.152]	[0.125]	[0.072]	[0.13]	[0.184]	[0.203]	[0.067]	[0.102]
outdegree - popularity	-0.214	-0.095	0.036	-0.19	-0.147	-0.12	-0.007	-0.469	-0.098	-0.129
	[0.076]	[0.252]	[0.101]	[0.134]	[0.061]	[0.063]	[0.202]	[0.132]	[0.036]	[0.043]
outdegree - activity	-0.045	-0.039	0.234	-0.031	0.09	0.023	-0.109	-0.012	0.027	0.065
	[0.046]	[0.1]	[0.102]	[0.073]	[0.029]	[0.028]	[0.087]	[0.04]	[0.022]	[0.03]
sex alter	-0.07	0.438	0.337	-0.675	0.317	-0.144	-0.702	-1.711	-0.03	-0.348
	[0.205]	[0.389]	[0.475]	[0.394]	[0.206]	[0.318]	[0.612]	[0.555]	[0.232]	[0.235]
sex ego	-0.204	0.104	-0.186	-0.255	-0.011	-0.062	0.114	-0.082	0.477	0.132
	[0.248]	[0.408]	[0.436]	[0.345]	[0.211]	[0.213]	[0.357]	[0.244]	[0.218]	[0.209]
same sex	0.515	0.624	1.417	-0.059	0.911	0.092	0.767	0.124	0.673	0.887
	[0.181]	[0.45]	[0.558]	[0.275]	[0.208]	[0.166]	[0.287]	[0.276]	[0.192]	[0.288]
grade alter	-0.213	-0.002	-0.118	0.179	-0.049	-0.118	0.597	0.464	0.227	0.137
	[0.18]	[0.181]	[0.21]	[0.179]	[0.109]	[0.109]	[0.736]	[0.245]	[0.184]	[0.14]
grade ego	-0.483	-0.29	-0.457	0.033	-0.18	-0.071	-0.64	-0.128	0.163	0.128
	[0.249]	[0.206]	[0.377]	[0.159]	[0.12]	[0.068]	[0.525]	[0.152]	[0.149]	[0.101]
grade similarity	0.385	0.661	0.319	-0.129	0.563	0.94	0.553	0.711	-0.624	-1.043
	[0.357]	[0.626]	[0.603]	[0.542]	[0.645]	[0.347]	[0.423]	[0.476]	[0.334]	[0.561]
TG alter	-0.083	0.037	-0.194	-0.067	-0.064	0.219	0.701	-0.395	-0.087	-0.044
	[0.067]	[0.144]	[0.113]	[0.068]	[0.044]	[0.106]	[0.527]	[0.159]	[0.069]	[0.09]
TG ego	-0.187	0.152	0.031	-0.013	-0.06	0.057	-0.326	-0.176	-0.012	-0.027
	[0.182]	[0.126]	[0.105]	[0.063]	[0.041]	[0.054]	[0.277]	[0.09]	[0.059]	[0.059]
TG similarity	-0.329	0.492	-0.093	-0.888	-0.63	0.325	0.807	-0.074	-0.432	0.162
	[0.699]	[0.707]	[0.883]	[0.595]	[0.365]	[0.312]	[0.604]	[0.559]	[0.362]	[0.387]
popularity alter	0.089	0.134	0.28	0.289	0.221	0.121	0.455	0.579	0.256	0.267
	[0.076]	[0.193]	[0.168]	[0.134]	[0.076]	[0.078]	[0.242]	[0.173]	[0.093]	[0.102]
popularity ego	0.114	-0.133	0.246	-0.056	0.005	0.002	0.203	-0.05	0.096	-0.009
	[0.108]	[0.185]	[0.191]	[0.096]	[0.068]	[0.045]	[0.213]	[0.066]	[0.056]	[0.068]
popularity ego x popularity alter	0.014	0.202	-0.16	0.054	0.034	0.036	-0.042	0.032	-0.022	0.053
	[0.013]	[0.161]	[0.074]	[0.037]	[0.03]	[0.017]	[0.06]	[0.023]	[0.021]	[0.033]
Overall maximum convergence ratio:	0.141	0.136	0.140	0.122	0.159	0.197	0.118	0.118	0.132	0.179

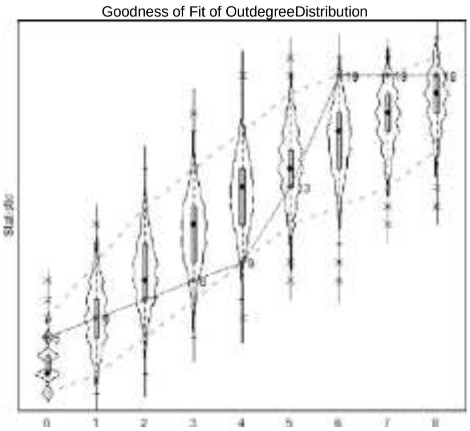
	11	12	13	14	15	16	17	18	19	20
outdegree (density)	-2.94	-2.577	-1.619	-1.298	-0.529	-1.03	0.047	-0.613	-1.105	-2.383
	[0.825]	[0.742]	[0.781]	[0.694]	[0.958]	[0.614]	[1.979]	[1.554]	[1.885]	[2.853]
reciprocity	0.682	1.709	1.182	1.263	0.998	0.535	1.538	2.236	1.649	1.722
	[0.603]	[0.687]	[0.582]	[0.43]	[0.472]	[0.389]	[0.815]	[1.165]	[0.997]	[1.272]
transitive triplets	0.075	0.12	0.366	0.272	0.22	0.081	0.451	1.439	0.365	-0.046
	[0.153]	[0.089]	[0.174]	[0.119]	[0.072]	[0.086]	[0.255]	[0.612]	[0.3]	[0.462]
transitive recipr. triplets	0.084	-0.124	-0.251	-0.267	-0.202	0.159	0.007	-1.049	-0.447	-0.007
	[0.2]	[0.142]	[0.27]	[0.182]	[0.08]	[0.117]	[0.373]	[0.905]	[0.504]	[0.524]
indegree - popularity	-0.073	0.065	0.071	-0.093	-0.2	-0.028	-0.317	-0.036	-0.331	0.298
	[0.129]	[0.065]	[0.077]	[0.1]	[0.129]	[0.054]	[0.387]	[0.159]	[0.302]	[0.256]
outdegree - popularity	-0.037	-0.089	-0.157	-0.038	0.09	-0.12	0.824	0.306	-0.122	0
	[0.076]	[0.041]	[0.126]	[0.086]	[0.064]	[0.04]	[1.005]	[0.474]	[0.126]	[NA]
outdegree - activity	0.105	0.077	-0.002	0.004	0.012	0.014	-0.618	-0.78	0.125	-0.157
	[0.036]	[0.027]	[0.045]	[0.036]	[0.023]	[0.026]	[0.317]	[0.37]	[0.076]	[0.291]
sex alter	0.586	0.12	0.867	-0.017	-0.297	0.301	0.533	0.127	0.203	0.133
	[0.363]	[0.231]	[0.375]	[0.233]	[0.289]	[0.212]	[0.706]	[0.655]	[0.874]	[0.722]
sex ego	-0.201	0.309	0.681	0.217	-0.304	0.046	-0.801	-0.786	-0.524	-1.268
	[0.244]	[0.237]	[0.402]	[0.238]	[0.168]	[0.2]	[0.733]	[0.964]	[0.719]	[1.077]
same sex	1.278	0.438	0.256	0.876	0.238	0.409	0.618	1.01	0.68	1.869
	[0.314]	[0.215]	[0.391]	[0.243]	[0.158]	[0.179]	[0.366]	[0.511]	[0.567]	[1.02]
grade alter	-0.136	-0.177	-0.016	0.361	0.033	0.182	0.266	0.052	0.149	0.195
	[0.181]	[0.192]	[0.129]	[0.183]	[0.117]	[0.121]	[0.305]	[0.346]	[0.374]	[0.302]
grade ego	-0.213	0.394	0.004	0.026	-0.168	0.141	0.085	-0.191	0.237	0.179
	[0.147]	[0.215]	[0.147]	[0.128]	[0.09]	[0.121]	[0.321]	[0.458]	[0.276]	[0.422]
grade similarity	1.078	-0.249	1.338	-0.293	0.021	-0.103	-0.172	1.186	-1.537	0.512
	[0.585]	[0.4]	[0.681]	[0.477]	[0.37]	[0.39]	[0.962]	[0.896]	[1.129]	[1.121]
TG alter	-0.013	0.061	-0.184	-0.086	0.059	0.083	-0.002	-0.184	0.073	0.2
	[0.059]	[0.048]	[0.089]	[0.074]	[0.062]	[0.048]	[0.17]	[0.168]	[0.127]	[0.192]
TG ego	-0.03	-0.005	-0.094	-0.082	-0.047	-0.002	-0.081	-0.176	-0.001	-0.083
	[0.056]	[0.05]	[0.085]	[0.072]	[0.044]	[0.05]	[0.188]	[0.168]	[0.091]	[0.19]
TG similarity	-0.302	-0.353	0.279	0.129	-0.968	0.364	0.141	-2.505	-0.682	-0.588
	[0.653]	[0.385]	[0.6]	[0.378]	[0.601]	[0.268]	[0.879]	[1.834]	[0.917]	[0.919]
popularity alter	0.456	0.21	0.027	0.126	0.276	0.288	0.336	0.044	0.484	0.374
	[0.166]	[0.082]	[0.113]	[0.082]	[0.179]	[0.099]	[0.348]	[0.188]	[0.232]	[0.239]
popularity ego	-0.032	-0.012	-0.159	0.011	0.015	-0.101	-0.094	-0.206	-0.233	-0.162
	[0.098]	[0.056]	[0.138]	[0.062]	[0.054]	[0.064]	[0.21]	[0.274]	[0.133]	[0.261]
popularity ego x popularity alter	-0.021	0.025	0.077	0.028	0.059	0.027	0.044	0.012	0.06	0.118
	[0.027]	[0.018]	[0.052]	[0.023]	[0.025]	[0.032]	[0.052]	[0.086]	[0.048]	[0.076]
Overall maximum convergence ratio:	0.162	0.193	0.191	0.172	0.196	0.185	0.137	0.141	0.158	0.097

Table 17: Results of the meta-analysis for friendship network and TG, Model 2

	est	se	N	p	tau2	Q	Qp
outdegree (density)	-1.227	0.309	20.000	0.000	0.843	33.660	0.020
reciprocity	1.080	0.128	20.000	0.000	0.000	9.396	0.966
transitive triplets	0.175	0.028	20.000	0.000	0.001	15.099	0.716
transitive recipr. triplets	-0.050	0.048	20.000	0.291	0.084	17.168	0.579
indegree - popularity	-0.026	0.022	20.000	0.247	0.018	24.336	0.184
outdegree - popularity	-0.101	0.017	19.000	0.000	0.020	23.785	0.162
outdegree - activity	0.032	0.011	20.000	0.004	0.026	32.112	0.030
sex alter	0.025	0.089	20.000	0.781	0.214	31.744	0.033
sex ego	0.011	0.068	20.000	0.873	0.098	20.187	0.383
same sex	0.563	0.085	20.000	0.000	0.245	35.240	0.013
grade alter	0.046	0.039	20.000	0.238	0.030	19.216	0.443
grade ego	-0.032	0.036	20.000	0.382	0.052	21.320	0.319
grade similarity	0.210	0.155	20.000	0.177	0.421	32.345	0.029
TG alter	0.012	0.024	20.000	0.606	0.001	20.605	0.359
TG ego	-0.041	0.017	20.000	0.015	0.000	14.739	0.739
TG similarity	0.009	0.112	20.000	0.933	0.000	15.070	0.718
popularity alter	0.206	0.026	20.000	0.000	0.000	19.179	0.445
popularity ego	-0.009	0.017	20.000	0.616	0.000	15.354	0.700
popularity ego x popularity alter	0.024	0.007	20.000	0.000	0.010	24.138	0.191
int. TG ego x TG similarity	-0.129	0.067	20.000	0.054	0.000	12.355	0.870

Table 18: Goodness of fit statistics, Friendship and TG, Model 2

	Indegree distribution	Outdegree distribution	Geodesic Distance	Triad Census
1	0.857	0.775	0.956	0.998
2	0.863	0.391	0.426	0.864
3	0.755	0.654	0.211	0.545
4	0.715	0.541	0.879	0.994
5	0.531	0.068	0.311	0.207
6	0.739	0.798	0.344	0.224
7	0.191	0.714	0.745	0.838
8	0.454	0.796	0.989	0.976
9	0.822	0.923	0.589	0.940
10	0.858	0.719	0.570	0.997
11	0.458	0.002	0.376	0.675
12	0.286	0.584	0.706	0.788
13	0.225	0.225	0.168	0.746
14	0.900	0.800	0.964	0.854
15	0.936	0.969	0.993	0.872
16	0.840	0.468	0.353	0.250
17	0.633	0.785	0.988	0.994
18	0.989	0.586	0.897	0.948
19	0.983	0.561	0.559	0.646
20	0.703	0.756	0.953	0.635



p: 0.002

(a) Class 11 Figure 6: Friendship and TG, Model 2

Table 19: Results of separate SAOMs for friendship network and TG, Model 2

	1	2	3	4	5	6	7	8	9	10
outdegree (density)	-0.702	-1.177	-4.713	1.078	-2.84	0.254	0.101	3.845	-1.09	-1.561
	[0.915]	[1.702]	[1.759]	[1.257]	[0.786]	[1.202]	[0.977]	[1.668]	[0.63]	[0.972]
reciprocity	1.131	0.951	1.654	0.535	1.639	0.604	0.751	1.155	1.06	1.678
	[0.413]	[0.807]	[0.946]	[0.578]	[0.587]	[0.463]	[0.655]	[0.437]	[0.395]	[0.65]
transitive triplets	0.232	0.225	0.135	0.046	-0.023	0.119	0.277	0.31	0.19	0.271
	[0.122]	[0.264]	[0.226]	[0.168]	[0.107]	[0.13]	[0.161]	[0.132]	[0.077]	[0.125]
transitive recipr. triplets	0.052	-0.094	-0.198	0.202	0.231	-0.005	-0.029	-0.092	-0.095	-0.225
	[0.152]	[0.309]	[0.301]	[0.209]	[0.177]	[0.196]	[0.242]	[0.204]	[0.139]	[0.194]
indegree - popularity	0.007	0.104	0.043	-0.121	0.067	-0.163	-0.232	-0.566	-0.094	-0.107
	[0.049]	[0.125]	[0.152]	[0.127]	[0.079]	[0.131]	[0.167]	[0.203]	[0.067]	[0.099]
outdegree - popularity	-0.225	-0.088	0.019	-0.188	-0.142	-0.117	-0.014	-0.463	-0.099	-0.122
	[0.089]	[0.242]	[0.108]	[0.129]	[0.056]	[0.067]	[0.16]	[0.127]	[0.037]	[0.037]
outdegree - activity	-0.045	-0.045	0.22	-0.033	0.089	0.026	-0.108	-0.013	0.026	0.062
	[0.063]	[0.107]	[0.115]	[0.07]	[0.028]	[0.032]	[0.073]	[0.038]	[0.021]	[0.029]
sex alter	-0.067	0.443	0.324	-0.675	0.296	-0.125	-0.722	-1.742	-0.032	-0.348
	[0.205]	[0.414]	[0.464]	[0.405]	[0.195]	[0.319]	[0.578]	[0.583]	[0.233]	[0.235]
sex ego	-0.233	0.102	-0.177	-0.299	-0.047	-0.052	0.143	-0.091	0.484	0.14
	[0.305]	[0.399]	[0.447]	[0.375]	[0.219]	[0.21]	[0.346]	[0.257]	[0.227]	[0.203]
same sex	0.52	0.592	1.398	-0.058	0.922	0.094	0.757	0.108	0.673	0.883
to	[0.177]	[0.464]	[0.594]	[0.281]	[0.229]	[0.18]	[0.274]	[0.284]	[0.194]	[0.28]
grade alter	-0.223	-0.002	-0.147	0.175	-0.049	-0.116	0.574	0.458	0.224	0.15
	[0.184]	[0.178]	[0.219]	[0.181]	[0.115]	[0.116]	[0.592]	[0.235]	[0.18]	[0.142]
grade ego	-0.458	-0.291	-0.449	0.036	-0.166	-0.072	-0.641	-0.132	0.157	0.122
	[0.344]	[0.208]	[0.477]	[0.161]	[0.117]	[0.067]	[0.459]	[0.149]	[0.157]	[0.103]
grade similarity	0.391	0.669	0.41	-0.097	0.562	0.954	0.577	0.736	-0.624	-1.143
	[0.36]	[0.646]	[0.689]	[0.538]	[0.646]	[0.347]	[0.457]	[0.475]	[0.324]	[0.617]
TG alter	-0.16	0.006	-0.063	-0.058	0.022	0.201	0.617	-0.413	-0.098	0.003
	[0.273]	[0.162]	[0.21]	[0.08]	[0.062]	[0.136]	[0.429]	[0.183]	[0.085]	[0.111]
TG ego	-0.164	0.164	0.001	-0.023	-0.087	0.062	-0.269	-0.168	-0.008	-0.046
	[0.242]	[0.138]	[0.103]	[0.076]	[0.046]	[0.059]	[0.304]	[0.096]	[0.062]	[0.069]
TG similarity	-0.691	0.448	0.089	-0.846	-0.231	0.31	0.911	-0.095	-0.441	0.239
	[1.113]	[0.693]	[0.897]	[0.611]	[0.404]	[0.327]	[0.631]	[0.565]	[0.371]	[0.408]
popularity alter	0.107	0.132	0.272	0.288	0.225	0.12	0.447	0.574	0.263	0.27
	[0.082]	[0.196]	[0.163]	[0.135]	[0.076]	[0.071]	[0.226]	[0.182]	[0.095]	[0.104]
popularity ego	0.108	-0.141	0.232	-0.056	0.014	0.003	0.203	-0.052	0.096	0.006
	[0.129]	[0.201]	[0.208]	[0.098]	[0.063]	[0.042]	[0.202]	[0.067]	[0.056]	[0.067]
popularity ego x popularity alter	0.012	0.202	-0.151	0.053	0.03	0.036	-0.043	0.034	-0.021	0.057
	[0.014]	[0.168]	[0.077]	[0.038]	[0.03]	[0.017]	[0.06]	[0.023]	[0.021]	[0.031]
int. TG ego x TG similarity	0.428	0.145	-0.703	-0.064	-0.285	0.059	0.452	0.137	0.048	-0.217
	[1.28]	[0.432]	[1.194]	[0.257]	[0.178]	[0.254]	[0.602]	[0.36]	[0.243]	[0.33]
Overall maximum convergence ratio:	0.224	0.114	0.124	0.127	0.153	0.205	0.166	0.175	0.112	0.221

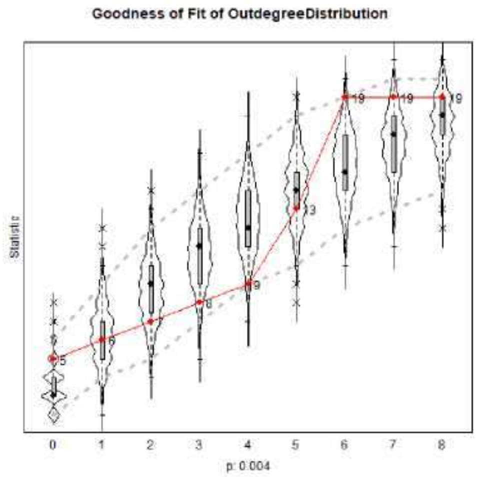
	11	12	13	14	15	16	17	18	19	20
outdegree (density)	-3.091	-2.604	-1.745	-1.274	-0.531	-1.046	0.094	-0.832	-1.163	-2.376
	[0.892]	[0.753]	[0.721]	[0.702]	[1.062]	[0.62]	[2.41]	[2.363]	[1.795]	[2.749]
reciprocity	0.704	1.629	1.179	1.291	1.072	0.523	1.583	2.288	1.693	1.929
	[0.712]	[0.627]	[0.609]	[0.415]	[0.586]	[0.388]	[0.904]	[1.631]	[1.012]	[1.358]
transitive triplets	0.086	0.112	0.362	0.282	0.234	0.077	0.458	1.454	0.357	-0.015
	[0.156]	[0.081]	[0.177]	[0.126]	[0.079]	[0.084]	[0.278]	[0.856]	[0.295]	[0.432]
transitive recipr. triplets	0.096	-0.107	-0.256	-0.287	-0.217	0.169	0.006	-1.103	-0.447	-0.031
	[0.214]	[0.135]	[0.284]	[0.18]	[0.096]	[0.1]	[0.375]	[1.098]	[0.501]	[0.572]
indegree - popularity	-0.086	0.066	0.076	-0.105	-0.195	-0.027	-0.308	-0.019	-0.341	0.298
	[0.151]	[0.064]	[0.08]	[0.1]	[0.137]	[0.056]	[0.498]	[0.159]	[0.291]	[0.232]
outdegree - popularity	-0.035	-0.087	-0.138	-0.031	0.083	-0.121	0.847	0.297	-0.119	0
	[0.088]	[0.043]	[0.117]	[0.084]	[0.063]	[0.042]	[1.639]	[0.774]	[0.122]	[NA]
outdegree - activity	0.105	0.08		0.003	0.008	0.014	-0.643	-0.817	0.131	-0.177
	[0.039]	[0.027]	0 [0.044]	[0.037]	[0.023]	[0.027]	[0.478]	[0.56]	[0.076]	[0.29]
sex alter	0.629	0.154	0.862	-0.025	-0.297	0.3	0.435	0.169	0.21	0.152
	[0.396]	[0.231]	[0.372]	[0.228]	[0.25]	[0.209]	[0.751]	[0.645]	[0.877]	[0.686]
sex ego	-0.19	0.341	0.703	0.218	-0.294	0.042	-0.768	-0.971	-0.557	-1.324
	[0.302]	[0.249]	[0.428]	[0.24]	[0.168]	[0.193]	[0.739]	[0.857]	[0.683]	[1.12]
same sex	1.274	0.438	0.279	0.884	0.256	0.409	0.691	1.04	0.74	1.842
	[0.325]	[0.231]	[0.388]	[0.239]	[0.155]	[0.173]	[0.421]	[0.54]	[0.584]	[0.879]
grade alter	-0.139	-0.167	-0.011	0.381	0.037	0.182	0.29	0.059	0.175	0.199
	[0.187]	[0.192]	[0.13]	[0.193]	[0.117]	[0.12]	[0.318]	[0.33]	[0.369]	[0.291]
grade ego	-0.167	0.384	0.022	0.027	-0.148	0.142	0.125	-0.378	0.13	0.154
	[0.151]	[0.217]	[0.148]	[0.128]	[0.093]	[0.121]	[0.314]	[0.488]	[0.304]	[0.414]
grade similarity	1.156	-0.267	1.497	-0.29	-0.012	-0.106	-0.114	1.406	-1.673	0.583
	[0.617]	[0.415]	[0.687]	[0.497]	[0.394]	[0.387]	[1.032]	[1.177]	[1.163]	[1.165]
TG alter	0.076	0.113	-0.069	-0.083	0.09	0.095	-0.097	-0.079	-0.065	0.355
	[0.085]	[0.085]	[0.1]	[0.108]	[0.069]	[0.082]	[0.146]	[0.141]	[0.154]	[0.28]
TG ego	-0.078	-0.016	-0.154	-0.082	-0.075	-0.004	-0.026	-0.368	0.073	-0.137
	[0.068]	[0.052]	[0.096]	[0.073]	[0.05]	[0.052]	[0.165]	[0.316]	[0.117]	[0.218]
TG similarity	0.409	-0.283	0.241	0.129	-0.688	0.379	-0.082	-2.307	-0.712	-0.69
	[0.776]	[0.378]	[0.6]	[0.385]	[0.589]	[0.271]	[0.966]	[1.851]	[0.899]	[0.862]
popularity alter	0.475	0.21	0.022	0.132	0.261	0.291	0.322	0.042	0.508	0.4
	[0.199]	[0.075]	[0.107]	[0.085]	[0.178]	[0.106]	[0.405]	[0.189]	[0.227]	[0.239]
popularity ego	-0.054	-0.018	-0.182	0.015	0.002	-0.106	-0.113	-0.201	-0.185	-0.159
	[0.106]	[0.051]	[0.134]	[0.061]	[0.055]	[0.061]	[0.226]	[0.261]	[0.136]	[0.3]
popularity ego x popularity alter	-0.026	0.024	0.069	0.028	0.058	0.027	0.038	0.01	0.065	0.118
	[0.029]	[0.019]	[0.048]	[0.023]	[0.023]	[0.032]	[0.067]	[0.081]	[0.047]	[0.068]
int. TG ego x TG similarity	-0.323	-0.205	-0.627	-0.021	-0.218	-0.043	0.594	-0.621	0.674	-0.638
	[0.18]	[0.279]	[0.441]	[0.336]	[0.201]	[0.218]	[0.586]	[0.592]	[0.519]	[0.721]
Overall maximum convergence ratio:	0.215	0.190	0.132	0.152	0.204	0.109	0.192	0.164	0.160	0.148

Table 20: Results of the meta-analysis for friendship network and TGB, Model 1

	est	se	N	p	tau2	Q	Qp
outdegree (density)	-1.190	0.2708	20	0.000	0.636	27.515	0.093
reciprocity	1.059	0.1260	20	0.000	0.000	10.405	0.942
transitive triplets	0.170	0.0290	20	0.000	0.001	15.244	0.707
transitive recipr. triplets	-0.049	0.0455	20	0.282	0.065	16.058	0.653
indegree - popularity	-0.030	0.0244	20	0.221	0.042	23.739	0.206
outdegree - popularity	-0.101	0.0216	19	0.000	0.053	26.077	0.098
outdegree - activity	0.029	0.0117	20	0.014	0.028	32.412	0.028
sex alter	0.038	0.0911	20	0.677	0.230	29.591	0.057
sex ego	0.014	0.0676	20	0.839	0.092	18.618	0.482
same sex	0.537	0.0838	20	0.000	0.233	33.674	0.020
grade alter	0.040	0.0384	20	0.298	0.000	17.345	0.566
grade ego	-0.022	0.0426	20	0.600	0.089	24.403	0.181
grade similarity	0.227	0.1504	20	0.131	0.400	30.754	0.043
TGB alter	0.003	0.0192	20	0.865	0.000	13.028	0.837
TGB ego	-0.036	0.0187	20	0.056	0.002	15.423	0.695
TGB similarity	-0.009	0.1895	20	0.961	0.369	23.582	0.213
popularity alter	0.194	0.0298	20	0.000	0.055	23.502	0.216
popularity ego	-0.019	0.0183	20	0.295	0.000	17.635	0.547
popularity ego x popularity alter	0.026	0.0072	20	0.000	0.012	25.011	0.160

Table 21: Goodness of fit statistics, Friendship and TGB, Model 1

	Indegree distribution	Outdegree distribution	Geodesic Distance	Triad Census
1	0.828	0.580	0.953	0.996
2	0.863	0.385	0.416	0.904
3	0.857	0.448	0.171	0.470
4	0.480	0.596	0.838	0.981
5	0.453	0.074	0.324	0.147
6	0.761	0.744	0.370	0.241
7	0.160	0.657	0.799	0.885
8	0.496	0.730	0.986	0.969
9	0.903	0.912	0.583	0.869
10	0.835	0.690	0.471	0.985
11	0.398	0.002	0.367	0.723
12	0.147	0.607	0.731	0.883
13	0.226	0.269	0.173	0.723
14	0.872	0.827	0.887	0.891
15	0.933	0.973	0.980	0.849
16	0.855	0.380	0.296	0.311
17	0.646	0.781	0.993	0.989
18	0.992	0.547	0.892	0.939
19	0.986	0.611	0.594	0.830
20	0.934	0.763	0.970	0.804



(a) Class 11

Figure 7: Classes with poor fit, Friendship and TGB, Model 1

Table 22: Results of separate SAOMs for friendship network and TGB, Model 1

	1	2	3	4	5	6	7	8	9	10
outdegree (density)	-0.907	-1.354	-7.465	0.865	-2.565	0.048	0.362	1.758	-1.169	-1.597
	[0.711]	[1.716]	[3.83]	[1.232]	[0.811]	[0.939]	[1.148]	[1.272]	[0.696]	[1.074]
reciprocity	1.125	0.956	2.092	0.552	1.59	0.6	0.742	1.109	1.077	1.743
	[0.378]	[0.831]	[1.237]	[0.59]	[0.539]	[0.455]	[0.767]	[0.407]	[0.465]	[0.65]
transitive triplets	0.218	0.217	0.117	0.004	-0.024	0.13	0.317	0.297	0.209	0.257
	[0.098]	[0.252]	[0.273]	[0.17]	[0.118]	[0.129]	[0.171]	[0.128]	[0.106]	[0.116]
transitive recipr. triplets	0.06	-0.069	-0.329	0.214	0.261	-0.017	-0.082	-0.094	-0.106	-0.205
	[0.144]	[0.298]	[0.398]	[0.213]	[0.2]	[0.205]	[0.306]	[0.201]	[0.175]	[0.163]
indegree - popularity	0.014	0.106	0.273	-0.163	0.061	-0.109	-0.174	-0.365	-0.085	-0.099
	[0.046]	[0.135]	[0.243]	[0.149]	[0.079]	[0.101]	[0.108]	[0.158]	[0.07]	[0.108]
outdegree - popularity	-0.209	-0.108	0.062	-0.119	-0.157	-0.12	-0.149	-0.303	-0.104	-0.129
	[0.074]	[0.221]	[0.117]	[0.12]	[0.052]	[0.065]	[0.112]	[0.093]	[0.039]	[0.044]
outdegree - activity	-0.035	-0.023	0.361	-0.007	0.085	0.012	-0.099	0.002	0.013	0.065
	[0.039]	[0.104]	[0.226]	[0.065]	[0.029]	[0.033]	[0.066]	[0.035]	[0.029]	[0.032]
sex alter	0.038	0.536	0.571	-0.521	0.284	0.062	-0.418	-1.039	-0.081	-0.268
	[0.187]	[0.405]	[0.728]	[0.386]	[0.203]	[0.289]	[0.476]	[0.37]	[0.233]	[0.239]
sex ego	-0.086	0.18	-1.138	-0.233	0.044	-0.135	0.163	0.044	0.442	0.197
	[0.256]	[0.412]	[1.454]	[0.345]	[0.24]	[0.209]	[0.435]	[0.229]	[0.325]	[0.182]
same sex	0.479	0.531	2.102	-0.07	0.901	0.158	0.679	0.107	0.627	0.844
	[0.185]	[0.528]	[1.133]	[0.273]	[0.232]	[0.171]	[0.272]	[0.268]	[0.208]	[0.27]
grade alter	-0.152	-0.026	0.007	0.158	-0.018	0.033	-0.259	0.511	0.187	0.134
	[0.18]	[0.192]	[0.208]	[0.189]	[0.118]	[0.126]	[0.319]	[0.212]	[0.205]	[0.138]
grade ego	-0.437	-0.325	-0.238	0.056	-0.123	0.108	-0.399	-0.01	0.13	0.135
	[0.209]	[0.207]	[0.691]	[0.158]	[0.12]	[0.098]	[0.307]	[0.141]	[0.192]	[0.093]
grade similarity	0.386	0.593	0.445	-0.098	0.264	1.292	0.515	0.8	-0.625	-0.89
	[0.346]	[0.65]	[0.679]	[0.501]	[0.644]	[0.403]	[0.423]	[0.445]	[0.434]	[0.527]
TGB alter	-0.035	0.099	0.065	-0.142	0.013	0.013	-0.14	-0.021	0.075	-0.044
	[0.054]	[0.191]	[0.12]	[0.137]	[0.043]	[0.104]	[0.127]	[0.073]	[0.121]	[0.099]
TGB ego	-0.132	0.239	0.629	0.007	-0.065	0.007	0.003	-0.039	-0.272	-0.087
	[0.076]	[0.22]	[0.539]	[0.127]	[0.041]	[0.085]	[0.158]	[0.057]	[0.351]	[0.08]
TGB similarity	-0.754	0.293	2.621	-0.556	-0.91	-1.695	0.068	0.852	1.593	-0.233
	[0.569]	[0.646]	[2.232]	[1.714]	[0.47]	[0.916]	[0.559]	[0.755]	[1.503]	[0.646]
popularity alter	0.093	0.119	0.037	0.359	0.24	0.076	0.575	0.399	0.206	0.271
	[0.074]	[0.184]	[0.183]	[0.162]	[0.093]	[0.067]	[0.217]	[0.149]	[0.085]	[0.109]
popularity ego	0.112	-0.143	0.525	-0.079	0.003	-0.036	0.17	-0.04	0.176	-0.023
	[0.083]	[0.188]	[0.54]	[0.098]	[0.06]	[0.049]	[0.152]	[0.061]	[0.119]	[0.064]
popularity ego x popularity alter	0.014	0.211	-0.153	0.057	0.04	0.034	-0.056	0.027	-0.026	0.053
	[0.013]	[0.174]	[0.095]	[0.037]	[0.031]	[0.017]	[0.056]	[0.022]	[0.022]	[0.03]
Overall maximum convergence ratio:	0.135	0.166	0.220	0.149	0.204	0.188	0.212	0.209	0.157	0.227

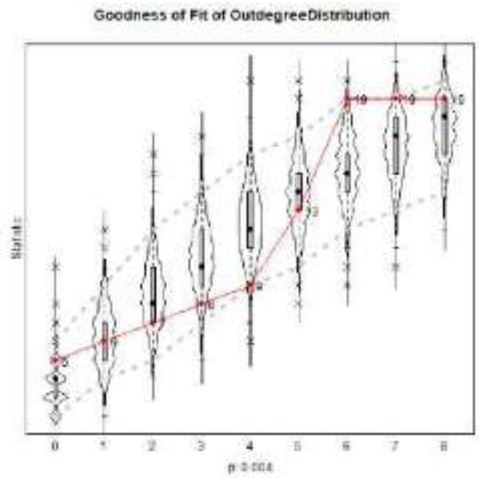
	11	12	13	14	15	16	17	18	19	20
outdegree (density)	-2.94	-2.685	-1.765	-1.248	-0.581	-1.171	0.134	-0.57	-0.699	-1.12
	[0.86]	[0.726]	[0.737]	[0.72]	[0.8]	[0.637]	[2.807]	[1.635]	[1.89]	[2.987]
reciprocity	0.705	1.526	1.191	1.134	0.858	0.513	1.5	2.254	1.939	2.699
	[0.687]	[0.594]	[0.581]	[0.398]	[0.465]	[0.393]	[1.045]	[1.263]	[1.055]	[1.857]
transitive triplets	0.061	0.101	0.337	0.242	0.203	0.064	0.433	1.508	0.409	0.117
	[0.151]	[0.082]	[0.201]	[0.12]	[0.074]	[0.088]	[0.299]	[0.731]	[0.293]	[0.453]
transitive recipr. triplets	0.087	-0.076	-0.239	-0.234	-0.176	0.168	0.057	-1.105	-0.543	0.071
	[0.202]	[0.122]	[0.297]	[0.175]	[0.085]	[0.119]	[0.427]	[0.917]	[0.504]	[0.726]
indegree - popularity	-0.094	0.067	0.099	-0.084	-0.176	-0.018	-0.35	-0.042	-0.422	0.177
	[0.149]	[0.067]	[0.074]	[0.098]	[0.099]	[0.056]	[0.833]	[0.181]	[0.316]	[0.234]
outdegree - popularity	-0.031	-0.071	-0.143	-0.058	0.082	-0.117	1.053	0.364	-0.143	0
	[0.084]	[0.043]	[0.108]	[0.075]	[0.052]	[0.037]	[2.949]	[0.537]	[0.131]	[NA]
outdegree - activity	0.112	0.078	-0.003	0.001	0.01	0.02	-0.734	-0.819	0.117	-0.346
	[0.037]	[0.025]	[0.048]	[0.034]	[0.022]	[0.026]	[0.902]	[0.485]	[0.075]	[0.329]
sex alter	0.555	-0.019	0.975	0.03	-0.343	0.306	0.478	0.033	0.682	0.346
	[0.412]	[0.253]	[0.378]	[0.228]	[0.253]	[0.212]	[0.739]	[0.717]	[0.884]	[0.824]
sex ego	-0.152	0.227	0.905	0.158	-0.319	0.101	-0.441	-0.826	-0.563	-1.878
	[0.269]	[0.241]	[0.441]	[0.236]	[0.153]	[0.207]	[0.799]	[1.1]	[0.802]	[1.393]
same sex	1.306	0.445	0.295	0.941	0.23	0.419	0.6	1.094	0.656	1.538
	[0.332]	[0.22]	[0.377]	[0.245]	[0.154]	[0.188]	[0.418]	[0.632]	[0.545]	[0.736]
grade alter	-0.243	-0.071	-0.046	0.236	-0.069	0.157	0.017	-0.182	0.36	0.561
	[0.194]	[0.194]	[0.134]	[0.155]	[0.107]	[0.122]	[0.525]	[0.379]	[0.451]	[0.445]
grade ego	-0.206	0.395	-0.046	-0.052	-0.152	0.152	0.397	-0.284	0.149	-0.264
	[0.147]	[0.215]	[0.159]	[0.114]	[0.09]	[0.138]	[0.506]	[0.582]	[0.297]	[0.496]
grade similarity	1.088	-0.265	1.03	-0.072	0.04	-0.169	-0.218	1.273	-2.113	0.747
	[0.579]	[0.382]	[0.638]	[0.463]	[0.39]	[0.371]	[1.203]	[0.995]	[1.242]	[1.136]
TGB alter	-0.03	0.09	-0.099	0.009	0.086	-0.007	-0.247	-0.162	-0.149	0.501
	[0.135]	[0.052]	[0.096]	[0.073]	[0.082]	[0.074]	[0.56]	[0.218]	[0.132]	[0.382]
TGB ego	0.055	0.031	-0.121	-0.027	-0.059	-0.11	0.305	-0.106	0.16	-0.622
	[0.124]	[0.046]	[0.099]	[0.081]	[0.053]	[0.102]	[0.348]	[0.24]	[0.126]	[0.508]
TGB similarity	0.564	0.759	0.141	1.504	-0.407	0.039	0.232	-3.336	1.593	-0.238
	[1.625]	[0.524]	[0.746]	[0.912]	[0.321]	[0.75]	[0.919]	[2.791]	[1]	[1.096]
popularity alter	0.487	0.202	-0.005	0.141	0.35	0.264	0.426	0.058	0.544	0.366
	[0.186]	[0.078]	[0.099]	[0.082]	[0.174]	[0.093]	[0.815]	[0.215]	[0.256]	[0.322]
popularity ego	-0.027	-0.004	-0.185	0.037	-0.005	-0.112	-0.191	-0.241	-0.255	-0.16
	[0.092]	[0.054]	[0.146]	[0.06]	[0.063]	[0.065]	[0.277]	[0.304]	[0.13]	[0.388]
popularity ego x popularity alter	-0.021	0.026	0.085	0.033	0.069	0.025	0.049	0.031	0.077	0.119
	[0.028]	[0.018]	[0.049]	[0.024]	[0.025]	[0.032]	[0.066]	[0.08]	[0.048]	[0.086]
Overall maximum convergence ratio:	0.189	0.210	0.151	0.166	0.148	0.115	0.179	0.217	0.124	0.189

Table 23: Results of the meta-analysis for friendship network and TGB, Model 2

	est	se	N	p	tau2	Q	Qp
outdegree (density)	-1.281	0.2837	20	0.000	0.679	28.194	0.080
reciprocity	1.143	0.1347	20	0.000	0.000	8.325	0.983
transitive triplets	0.183	0.0304	20	0.000	0.002	11.371	0.911
transitive recipr. triplets	-0.056	0.0463	20	0.227	0.057	14.928	0.727
indegree - popularity	-0.021	0.0245	20	0.395	0.040	23.327	0.223
outdegree - popularity	-0.107	0.0170	19	0.000	0.011	20.612	0.299
outdegree - activity	0.027	0.0121	20	0.026	0.029	31.209	0.038
sex alter	0.028	0.1044	20	0.789	0.286	31.356	0.037
sex ego	0.028	0.0800	20	0.726	0.171	25.605	0.142
same sex	0.541	0.0895	20	0.000	0.255	35.496	0.012
grade alter	0.040	0.0402	20	0.315	0.000	17.556	0.552
grade ego	-0.029	0.0448	20	0.521	0.088	23.292	0.225
grade similarity	0.214	0.1665	20	0.198	0.476	35.668	0.012
TGB alter	-0.003	0.0243	18	0.911	0.000	9.682	0.916
TGB ego	-0.033	0.0236	18	0.162	0.009	18.980	0.330
TGB similarity	-0.130	0.2151	18	0.546	0.410	18.169	0.378
popularity alter	0.181	0.0279	20	0.000	0.036	21.921	0.288
popularity ego	-0.016	0.0212	20	0.445	0.032	20.422	0.370
popularity ego x popularity alter	0.024	0.0070	20	0.001	0.010	23.798	0.204
int. TGB ego x TGB similarity	-0.092	0.1046	18	0.379	0.258	25.546	0.083

Table 24: Goodness of fit statistics, Friendship and TGB, Model 2

	Indegree distribution	Outdegree distribution	Geodesic Distance	Triad Census
1	0.844	0.534	0.963	0.996
2	0.874	0.341	0.417	0.910
3	0.866	0.375	0.185	0.540
4	0.526	0.622	0.854	0.983
5	0.468	0.087	0.352	0.247
6	0.773	0.738	0.374	0.158
7	0.125	0.666	0.790	0.839
8	0.501	0.760	0.985	0.955
9	0.886	0.941	0.557	0.951
10	0.855	0.752	0.500	0.987
11	0.567	0.004	0.481	0.643
12	0.152	0.589	0.745	0.897
13	0.130	0.229	0.131	0.713
14	0.894	0.790	0.944	0.879
15	0.928	0.984	0.974	0.834
16	0.816	0.474	0.256	0.109
17	0.611	0.736	0.989	0.992
18	0.990	0.515	0.916	0.956
19	0.986	0.616	0.552	0.784
20	0.892	0.832	0.974	0.808



(a) Class 11

Figure 8: Classes with poor fit, Friendship and TGB, Model 2

Table 25: Results of separate SAOMs for friendship network and TGB, Model 2

	1	2	3	4	5	6	7	8	9	10
outdegree (density)	-0.913	-1.422	-7.683	0.937	-2.437	0.195	0.281	1.766	-1.323	-1.778
	[0.647]	[1.66]	[4.211]	[1.309]	[0.735]	[1.267]	[1.088]	[1.256]	[0.573]	[0.923]
reciprocity	1.13	0.918	2.117	0.566	1.56	0.586	0.755	1.1	1.058	1.79
	[0.414]	[0.829]	[1.483]	[0.565]	[0.571]	[0.536]	[0.724]	[0.426]	[0.425]	[0.551]
transitive triplets	0.219	0.199	0.11	0.015	-0.01	0.126	0.316	0.296	0.18	0.279
	[0.103]	[0.247]	[0.281]	[0.171]	[0.121]	[0.147]	[0.16]	[0.123]	[0.087]	[0.142]
transitive recipr. triplets	0.06	-0.055	-0.339	0.196	0.24	-0.006	-0.077	-0.086	-0.094	-0.247
	[0.147]	[0.295]	[0.444]	[0.203]	[0.19]	[0.243]	[0.301]	[0.193]	[0.154]	[0.197]
indegree - popularity	0.016	0.114	0.284	-0.161	0.063	-0.107	-0.171	-0.374	-0.062	-0.088
	[0.044]	[0.136]	[0.236]	[0.16]	[0.071]	[0.121]	[0.115]	[0.154]	[0.058]	[0.09]
outdegree - popularity	-0.213	-0.103	0.068	-0.114	-0.16	-0.123	-0.152	-0.307	-0.094	-0.125
	[0.073]	[0.237]	[0.124]	[0.135]	[0.059]	[0.069]	[0.111]	[0.096]	[0.034]	[0.04]
outdegree - activity	-0.041	-0.017	0.371	-0.008	0.08	-0.001	-0.096	0.004	0.027	0.065
	[0.042]	[0.102]	[0.261]	[0.069]	[0.032]	[0.037]	[0.056]	[0.034]	[0.022]	[0.033]
sex alter	0.022	0.528	0.561	-0.508	0.299	0.032	-0.425	-1.055	-0.118	-0.263
	[0.193]	[0.425]	[0.738]	[0.387]	[0.211]	[0.31]	[0.481]	[0.35]	[0.208]	[0.226]
sex ego	-0.003	0.178	-1.186	-0.341	-0.017	-0.217	0.125	0.037	0.48	0.188
	[0.267]	[0.394]	[1.516]	[0.373]	[0.215]	[0.232]	[0.424]	[0.226]	[0.203]	[0.186]
same sex	0.472	0.526	2.146	-0.083	0.9	0.132	0.719	0.106	0.668	0.859
	[0.181]	[0.501]	[1.141]	[0.276]	[0.228]	[0.184]	[0.277]	[0.273]	[0.199]	[0.27]
grade alter	-0.175	-0.021	0.002	0.156	0.01	0.006	-0.26	0.53	0.194	0.136
	[0.19]	[0.188]	[0.215]	[0.185]	[0.113]	[0.116]	[0.339]	[0.215]	[0.172]	[0.143]
grade ego	-0.482	-0.327	-0.241	0.045	-0.085	0.133	-0.399	-0.007	0.144	0.147
	[0.235]	[0.217]	[0.636]	[0.166]	[0.122]	[0.107]	[0.301]	[0.132]	[0.15]	[0.107]
grade similarity	0.426	0.585	0.45	-0.15	0.021	1.364	0.545	0.774	-0.585	-0.89
	[0.366]	[0.651]	[0.677]	[0.514]	[0.671]	[0.412]	[0.426]	[0.455]	[0.317]	[0.559]
TGB alter	0.002	0.124	0.077	-0.178	-0.049	0.067	-0.088	-0.002	0	0.003
	[0.062]	[0.249]	[0.15]	[0.164]	[0.048]	[0.098]	[0.149]	[0.075]	[NA]	[0.113]
TGB ego	-0.212	0.238	0.641	0.103	0.013	-0.117	-0.041	-0.057	0	-0.114
	[0.127]	[0.229]	[0.554]	[0.169]	[0.051]	[0.132]	[0.134]	[0.063]	[NA]	[0.087]
TGB similarity	-0.597	0.309	2.736	-0.615	-1.315	-1.186	-0.09	0.895	0	-0.128
	[0.596]	[0.634]	[2.437]	[1.885]	[0.489]	[1.02]	[0.574]	[0.772]	[NA]	[0.648]
popularity alter	0.094	0.117	0.035	0.361	0.235	0.074	0.586	0.408	0.194	0.256
	[0.069]	[0.188]	[0.179]	[0.18]	[0.08]	[0.07]	[0.244]	[0.145]	[0.073]	[0.1]
popularity ego	0.134	-0.144	0.55	-0.101	0.014	-0.094	0.165	-0.042	0.091	-0.025
	[0.104]	[0.191]	[0.547]	[0.107]	[0.062]	[0.066]	[0.149]	[0.061]	[0.048]	[0.063]
popularity ego x popularity alter	0.015	0.213	-0.154	0.059	0.035	0.031	-0.057	0.027	-0.026	0.052
	[0.014]	[0.172]	[0.103]	[0.038]	[0.031]	[0.016]	[0.06]	[0.022]	[0.02]	[0.031]
int. TGB ego x TGB similarity	-0.28	-0.114	-0.09	0.316	0.424	-0.357	-0.487	-0.153	0	-0.213
	[0.279]	[0.746]	[0.615]	[0.355]	[0.19]	[0.242]	[0.482]	[0.265]	[NA]	[0.244]
Overall maximum convergence ratio:	0.139	0.121	0.145	0.176	0.150	0.146	0.163	0.163	0.086	0.234

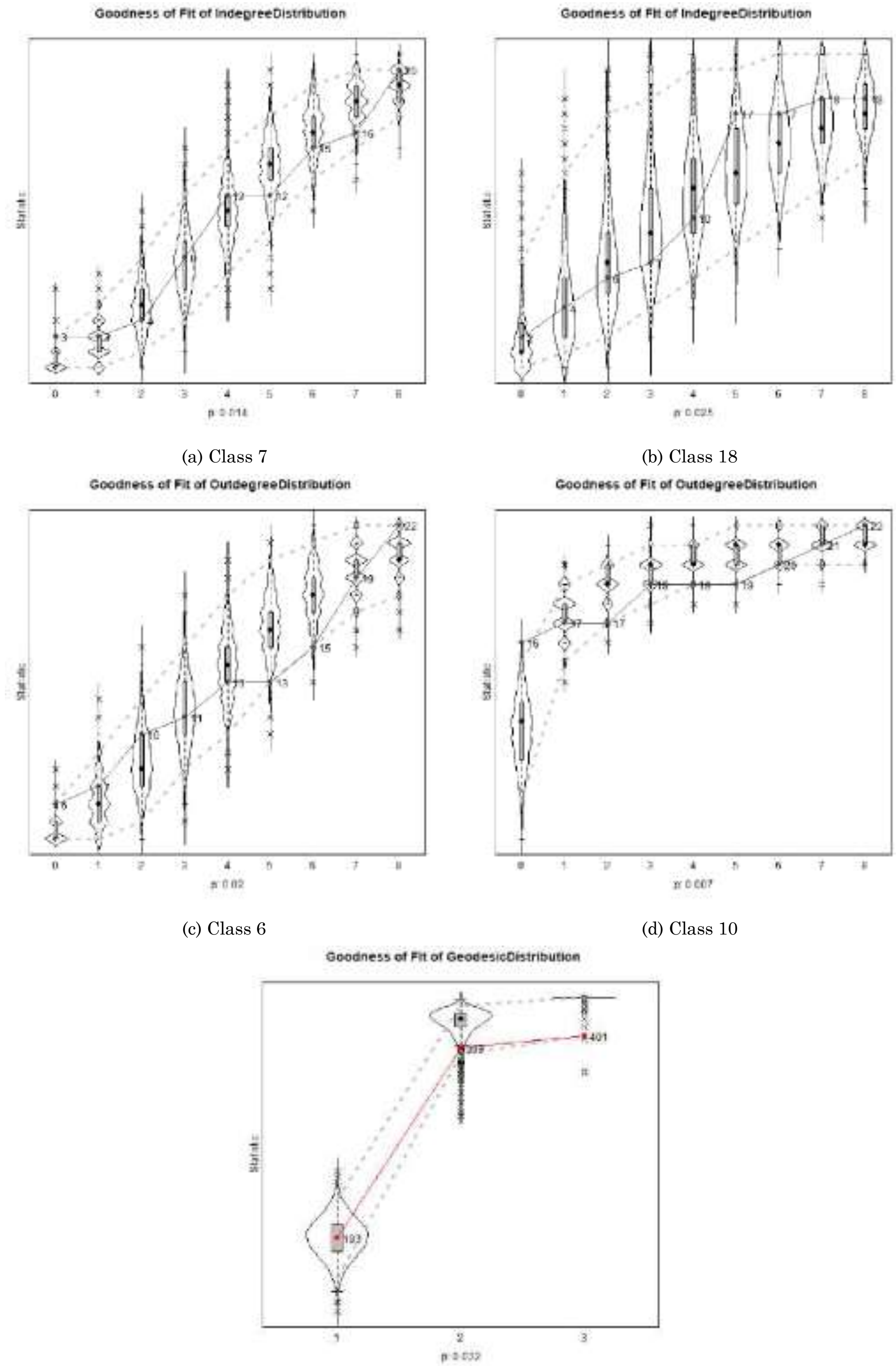
	11	12	13	14	15	16	17	18	19	20
outdegree (density)	-3.726	-2.633	-1.84	-1.291	-0.444	-1.518	0.093	-0.396	-0.639	-1.094
	[1.048]	[0.72]	[0.787]	[0.671]	[0.962]	[0.882]	[4.752]	[1.83]	[2.04]	[2.888]
reciprocity	0.578	1.52	1.235	1.293	0.896	1.102	1.519	2.52	1.922	2.845
	[0.768]	[0.621]	[0.63]	[0.442]	[0.507]	[0.519]	[1.753]	[2.01]	[1.027]	[1.644]
transitive triplets	0.043	0.102	0.333	0.279	0.219	0.131	0.432	1.695	0.393	0.112
	[0.199]	[0.08]	[0.177]	[0.13]	[0.076]	[0.134]	[0.522]	[1.196]	[0.287]	[0.374]
transitive recipr. triplets	0.16	-0.066	-0.216	-0.285	-0.187	0.193	0.061	-1.258	-0.476	0.059
	[0.273]	[0.121]	[0.277]	[0.189]	[0.086]	[0.155]	[0.465]	[1.455]	[0.482]	[0.627]
indegree - popularity	-0.126	0.065	0.103	-0.098	-0.196	-0.024	-0.348	-0.053	-0.418	0.167
	[0.163]	[0.066]	[0.069]	[0.108]	[0.144]	[0.083]	[1.168]	[0.174]	[0.346]	[0.224]
outdegree - popularity	-0.017	-0.075	-0.154	-0.022	0.091	-0.188	1.057	0.364	-0.155	0
	[0.101]	[0.044]	[0.114]	[0.097]	[0.076]	[0.095]	[3.817]	[0.552]	[0.129]	[NA]
outdegree - activity	0.13	0.077	-0.003	-0.004	0.008	0.022	-0.739	-1.009	0.119	-0.349
	[0.045]	[0.025]	[0.049]	[0.035]	[0.023]	[0.034]	[0.976]	[0.716]	[0.079]	[0.337]
sex alter	0.855	-0.019	0.994	-0.006	-0.393	0.5	0.5	0.121	0.668	0.31
	[0.445]	[0.245]	[0.389]	[0.225]	[0.324]	[0.428]	[1.497]	[0.667]	[0.893]	[0.783]
sex ego	-0.534	0.224	1.046	0.192	-0.341	0.146	-0.579	-1.177	-0.491	-2.126
	[0.362]	[0.244]	[0.473]	[0.236]	[0.167]	[0.303]	[1.39]	[1.069]	[0.829]	[1.572]
same sex	1.508	0.441	0.244	0.896	0.23	0.373	0.587	1.227	0.618	1.512
	[0.38]	[0.217]	[0.39]	[0.234]	[0.154]	[0.257]	[0.558]	[0.853]	[0.576]	[0.699]
grade alter	-0.311	-0.07	-0.046	0.305	-0.067	0.111	0.018	-0.154	0.398	0.592
	[0.24]	[0.205]	[0.149]	[0.173]	[0.109]	[0.199]	[0.538]	[0.458]	[0.464]	[0.463]
grade ego	-0.269	0.384	-0.087	-0.047	-0.152	0.225	0.372	-0.517	0.139	-0.229
	[0.181]	[0.216]	[0.17]	[0.12]	[0.086]	[0.519]	[0.53]	[0.666]	[0.311]	[0.511]
grade similarity	1.534	-0.267	1.164	-0.112	0.099	-0.544	-0.245	1.553	-2.239	0.708
	[0.717]	[0.391]	[0.661]	[0.455]	[0.394]	[0.588]	[1.07]	[1.254]	[1.311]	[1.18]
TGB alter	0.067	0.082	-0.031		-0.016	-0.622	-0.199	0.034	-0.213	0.671
	[0.195]	[0.063]	[0.127]	0 [NA]	[0.1]	[0.828]	[0.605]	[0.238]	[0.175]	[0.465]
TGB ego	-0.262	0.034	-0.242		-0.057	-0.444	0.229	-0.549	0.171	-0.727
	[0.167]	[0.054]	[0.131]	0 [NA]	[0.056]	[0.823]	[0.403]	[0.603]	[0.135]	[0.591]
TGB similarity	0.774	0.693	0.359		-0.529	0.894	0.346	-2.168	1.568	-0.322
	[2.094]	[0.582]	[0.821]	0 [NA]	[0.311]	[3.119]	[1.542]	[3.425]	[1.082]	[1.086]
popularity alter	0.55	0.199	-0.001	0.126	0.387	0.336	0.433	0.051	0.531	0.411
	[0.212]	[0.078]	[0.109]	[0.079]	[0.234]	[0.263]	[1.187]	[0.229]	[0.275]	[0.314]
popularity ego	-0.071	-0.007	-0.131	0.03	-0.035	-0.168	-0.188	-0.285	-0.272	-0.196
	[0.129]	[0.055]	[0.144]	[0.063]	[0.065]	[0.116]	[0.325]	[0.306]	[0.139]	[0.353]
popularity ego x popularity alter	-0.039	0.026	0.083	0.029	0.061	0.035	0.049	0.036	0.083	0.125
	[0.035]	[0.019]	[0.051]	[0.023]	[0.027]	[0.037]	[0.059]	[0.089]	[0.054]	[0.084]
int. TGB ego x TGB similarity	-0.752	0.032	-0.559	0 [NA]	0.403	5.232	-0.256	-1.214	0.255	-0.808
	[0.301]	[0.162]	[0.433]		[0.218]	[6.739]	[0.508]	[1.051]	[0.408]	[0.943]
Overall maximum convergence ratio:	0.148	0.190	0.186	0.140	0.248	0.224	0.159	0.237	0.167	0.173

Table 26: Results of the meta-analysis for negative network and PGG, Model 1

	est	se	N	p	tau2	Q	Qp
outdegree (density)	-2.121	0.147	20	0.000	0.202	18.790	0.470
reciprocity	0.583	0.081	20	0.000	0.002	18.129	0.514
transitive triplets	-0.099	0.026	20	0.000	0.052	22.214	0.274
indegree - popularity	0.115	0.014	20	0.000	0.001	16.158	0.647
outdegree - activity	0.096	0.009	20	0.000	0.017	22.467	0.262
indegree - activity	-0.021	0.043	5	0.626	0.000	2.176	0.703
outdegree-trunc(1)	-2.775	0.990	4	0.005	1.245	5.573	0.134
sex alter	0.017	0.082	20	0.837	0.191	31.816	0.033
sex ego	0.104	0.062	20	0.092	0.000	26.822	0.109
same sex	-0.511	0.058	20	0.000	0.001	26.792	0.110
grade alter	0.008	0.035	20	0.829	0.041	16.123	0.649
grade ego	0.036	0.041	20	0.375	0.090	24.820	0.167
grade similarity	-0.005	0.132	20	0.973	0.325	30.929	0.041
PGG alter	0.002	0.025	20	0.944	0.045	18.440	0.493
PGG ego	0.013	0.030	20	0.654	0.068	24.737	0.169
PGG similarity	-0.049	0.110	20	0.655	0.217	24.082	0.193
popularity alter	-0.134	0.021	20	0.000	0.002	25.447	0.146
popularity ego	-0.026	0.016	20	0.102	0.000	13.521	0.811
popularity ego x popularity alter	-0.015	0.005	20	0.002	0.000	13.987	0.784

Table 27: Goodness of fit statistics, Negative and PGG, Model 1

	Indegree distribution	Outdegree distribution	Geodesic Distance	Triad Census
1	0.716	0.396	0.922	0.732
2	0.975	0.807	0.952	0.835
3	0.626	0.974	1.000	0.663
4	0.195	0.322	0.194	0.090
5	0.333	0.160	0.065	0.970
6	0.969	0.020	0.020	0.997
7	0.014	0.777	0.149	0.737
8	0.936	0.488	0.071	0.970
9	0.877	0.425	0.080	0.904
10	0.852	0.007	0.769	0.780
11	0.677	0.831	0.883	0.711
12	0.188	0.545	0.869	0.990
13	0.988	0.108	0.037	0.513
14	0.394	0.216	0.594	0.947
15	0.981	0.118	0.284	0.291
16	0.692	0.692	0.078	0.757
17	0.985	0.833	0.971	1.000
18	0.025	0.403	0.344	0.522
19	0.983	0.800	0.426	0.706
20	0.440	0.543	0.946	0.746



(a) Class 7

(b) Class 18

(c) Class 6

(d) Class 10

(e) Class 13

Figure 9: Classes with poor fit, Negative and PGG, Model 1

Table 28: Results of separate SAOMs for negative network and PGG, Model 1

	1	2	3	4	5	6	7	8	9	10
outdegree (density)	-1.952	2.595	0.079	-2.931	-2.347	-1.236	-2.207	-1.771	-1.548	-1.837
	[0.347]	[13.541]	[1.668]	[0.861]	[0.58]	[0.795]	[0.641]	[0.367]	[0.397]	[5.697]
reciprocity	0.699	2.286	0.004	0.431	1.112	0.63	1.115	0.429	0.299	0.074
	[0.224]	[1.378]	[0.498]	[0.52]	[0.292]	[0.47]	[0.408]	[0.212]	[0.266]	[0.971]
transitive triplets	-0.057	1.16	-0.161	0.396	-0.203	-0.001	-0.104	-0.247	0.037	0.006
	[0.05]	[1.203]	[0.166]	[0.268]	[0.08]	[0.158]	[0.084]	[0.062]	[0.069]	[0.904]
indegree - popularity	0.093	-0.096	-0.313	0.071	0.091	0.12	0.089	0.11	0.034	-1.177
	[0.036]	[0.5]	[0.279]	[0.122]	[0.068]	[0.047]	[0.065]	[0.045]	[0.052]	[0.775]
outdegree - activity	0.072	-1.468	0.14	0.058	0.152	0.051	0.117	0.098	0.034	0.264
	[0.019]	[4.484]	[0.055]	[0.073]	[0.028]	[0.049]	[0.035]	[0.02]	[0.027]	[0.147]
indegree - activity	0	0	0	0	0	-0.166	0	0	0	0.112
	[NA]	[NA]	[NA]	[NA]	[NA]	[0.119]	[NA]	[NA]	[NA]	[3.613]
outdegree-trunc(1)	0	0	0	1.349	0	0	0	-2.499	0	0
	[NA]	[NA]	[NA]	[2.14]	[NA]	[NA]	[NA]	[1.287]	[NA]	[NA]
sex alter	-0.019	-1.156	-1.967	0.111	-0.485	-0.166	0.087	0.309	0.249	-0.52
	[0.156]	[1.41]	[1.025]	[0.45]	[0.411]	[0.228]	[0.239]	[0.19]	[0.24]	[1.047]
sex ego	-0.027	3.629	-1.057	0.982	-0.341	-0.615	-0.031	0.063	0.34	-0.351
	[0.146]	[13.238]	[0.537]	[0.549]	[0.385]	[0.5]	[0.251]	[0.165]	[0.237]	[0.917]
same sex	-0.531	-1.346	-0.619	0.085	-1.4	-0.149	-0.405	-0.663	-0.501	-1.066
	[0.148]	[1.044]	[0.48]	[0.387]	[0.424]	[0.197]	[0.211]	[0.164]	[0.204]	[0.948]
grade alter	0.089	-0.375	-0.038	0.075	0.139	0.081	0.312	-0.179	-0.272	0.075
	[0.135]	[0.619]	[0.227]	[0.224]	[0.176]	[0.104]	[0.206]	[0.094]	[0.15]	[0.472]
grade ego	0.201	-1.575	0.369	-0.221	0.223	0.103	0.518	-0.152	-0.203	-0.796
	[0.142]	[5.225]	[0.246]	[0.243]	[0.157]	[0.111]	[0.247]	[0.092]	[0.149]	[1.19]
grade similarity	-0.033	-0.42	0.495	-0.475	-0.537	-0.253	-0.231	-0.598	0.135	3.154
	[0.248]	[1.763]	[0.481]	[0.772]	[0.883]	[0.432]	[0.33]	[0.305]	[0.301]	[1.673]
PGG alter	0.046	-0.137	0.154	0.06	-0.007	0.09	0.006	-0.024	0.025	0.063
	[0.065]	[0.923]	[0.313]	[0.167]	[0.113]	[0.08]	[0.151]	[0.077]	[0.094]	[0.38]
PGG ego	-0.107	-6.237	0.445	0.144	-0.145	0.182	0.067	-0.02	-0.02	0.055
	[0.07]	[18.487]	[0.267]	[0.153]	[0.122]	[0.151]	[0.165]	[0.066]	[0.086]	[0.676]
PGG similarity	-0.139	0.918	0.702	1.753	-0.564	0.746	0.563	0.097	-0.385	0.957
	[0.285]	[1.926]	[0.731]	[0.858]	[0.432]	[0.427]	[0.527]	[0.363]	[0.304]	[2.957]
popularity alter	-0.098	-0.07	-0.811	-0.235	-0.263	-0.061	-0.343	-0.167	-0.139	-1.36
	[0.053]	[0.443]	[0.416]	[0.169]	[0.103]	[0.074]	[0.128]	[0.058]	[0.083]	[0.597]
popularity ego	0.016	-3.663	-0.267	0.072	-0.06	-0.248	-0.01	-0.03	-0.105	0.293
	[0.034]	[10.815]	[0.134]	[0.12]	[0.068]	[0.184]	[0.105]	[0.04]	[0.048]	[2.005]
popularity ego x popularity alter	-0.023	0.131	0.013	-0.114	0.019	-0.017	0.013	-0.016	0.004	0.05
	[0.008]	[0.367]	[0.031]	[0.057]	[0.034]	[0.02]	[0.034]	[0.015]	[0.017]	[0.095]
Overall maximum convergence ratio:	0.108	0.154	0.161	0.147	0.119	0.172	0.126	0.210	0.098	0.208

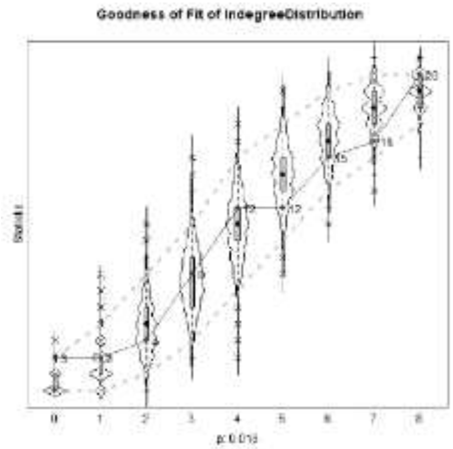
	11	12	13	14	15	16	17	18	19	20
outdegree (density)	-1.624	-2.96	-1.653	-2.968	-3.063	-2.566	-2.163	-2.328	-0.398	-1.749
	[0.696]	[0.6]	[0.835]	[0.492]	[0.534]	[1.067]	[0.838]	[0.401]	[1.252]	[0.784]
reciprocity	0.484	0.988	0.843	0.684	-0.102	0.427	0.453	0.672	-0.829	-0.038
	[0.36]	[0.345]	[0.326]	[0.27]	[0.413]	[0.303]	[0.895]	[0.446]	[1.203]	[0.547]
transitive triplets	-0.22	-0.053	0.006	-0.117	-0.241	-0.116	-0.744	-0.025	-0.336	-0.272
	[0.115]	[0.065]	[0.115]	[0.059]	[0.184]	[0.09]	[0.808]	[0.094]	[0.532]	[0.223]
indegree - popularity	0.112	0.11	0.094	0.168	0.189	0.12	-0.127	0.168	-0.102	0.032
	[0.073]	[0.068]	[0.058]	[0.041]	[0.056]	[0.053]	[0.201]	[0.042]	[0.25]	[0.187]
outdegree - activity	0.091	0.103	0.066	0.109	0.171	0.102	0.033	0.078	0.157	0.211
	[0.047]	[0.026]	[0.037]	[0.026]	[0.049]	[0.028]	[0.069]	[0.034]	[0.131]	[0.065]
indegree - activity	0	0	-0.064	0	0	0.013	0	0.017	0	0
	[NA]	[NA]	[0.108]	[NA]	[NA]	[0.108]	[NA]	[0.058]	[NA]	[NA]
outdegree-trunc(1)	-4.164	0	0	0	0	0	0	0	-4.174	0
	[1.333]	[NA]	[NA]	[NA]	[NA]	[NA]	[NA]	[NA]	[1.599]	[NA]
sex alter	-0.382	-0.291	-0.227	0.272	1.262	0.022	-0.509	0.371	-1.407	-0.68
	[0.251]	[0.24]	[0.334]	[0.184]	[0.42]	[0.233]	[0.737]	[0.228]	[1.209]	[0.504]
sex ego	-0.104	0.119	0.079	0.25	0.808	0.384	1.686	0.414	-0.837	-0.219
	[0.237]	[0.221]	[0.293]	[0.193]	[0.333]	[0.298]	[0.954]	[0.234]	[0.688]	[0.403]
same sex	-0.336	-0.314	-0.448	-0.427	-1.036	-0.5	0.43	-0.99	-1.59	-1.539
	[0.249]	[0.227]	[0.286]	[0.201]	[0.344]	[0.191]	[0.591]	[0.307]	[0.706]	[0.56]
grade alter	0.014	0.042	0.109	0.015	0.041	0.053	-0.007	-0.181	0.19	0.279
	[0.121]	[0.187]	[0.11]	[0.104]	[0.129]	[0.119]	[0.23]	[0.129]	[0.458]	[0.249]
grade ego	0.018	-0.095	-0.115	0.141	-0.035	0.241	-0.017	0.044	-0.078	0.111
	[0.129]	[0.173]	[0.097]	[0.114]	[0.115]	[0.118]	[0.193]	[0.145]	[0.231]	[0.206]
grade similarity	-0.692	-0.661	-0.154	0.794	1.425	0.734	0.98	0.057	-1.408	0.268
	[0.449]	[0.405]	[0.419]	[0.432]	[0.651]	[0.378]	[0.992]	[0.392]	[0.897]	[0.77]
PGG alter	0.166	-0.075	0.024	-0.041	-0.042	0.037	-0.404	-0.14	0.383	-0.219
	[0.077]	[0.09]	[0.085]	[0.072]	[0.098]	[0.068]	[0.262]	[0.059]	[0.329]	[0.289]
PGG ego	0.138	-0.168	0.107	-0.151	-0.001	0.18	0.013	0.022	0.157	-0.08
	[0.082]	[0.108]	[0.088]	[0.083]	[0.089]	[0.084]	[0.202]	[0.055]	[0.199]	[0.214]
PGG similarity	-0.456	-0.321	0.066	0.519	-0.459	-0.23	-0.656	-0.349	0.221	1.109
	[0.293]	[0.445]	[0.411]	[0.303]	[0.453]	[0.336]	[0.708]	[0.248]	[1.142]	[0.996]
popularity alter	-0.27	-0.116	-0.106	-0.05	-0.062	-0.224	-0.28	0.002	-0.718	-0.489
	[0.125]	[0.063]	[0.095]	[0.066]	[0.116]	[0.088]	[0.195]	[0.08]	[0.407]	[0.237]
popularity ego	-0.064	-0.001	0.06	-0.001	-0.02	0.1	-0.098	0.095	-0.003	-0.112
	[0.063]	[0.045]	[0.136]	[0.059]	[0.087]	[0.213]	[0.115]	[0.12]	[0.116]	[0.134]
popularity ego x popularity alter	-0.019	-0.006	-0.019	-0.036	-0.024	0.024	-0.033	-0.026	-0.028	-0.086
	[0.021]	[0.014]	[0.033]	[0.023]	[0.037]	[0.027]	[0.065]	[0.036]	[0.046]	[0.052]
Overall maximum convergence ratio:	0.136	0.133	0.126	0.120	0.134	0.132	0.110	0.136	0.214	0.141

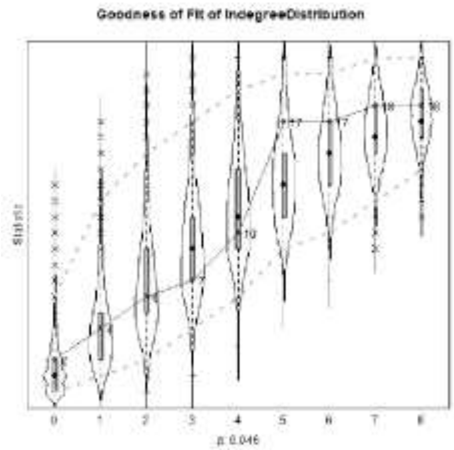
Table 29: Results of the meta-analysis for negative network and PGG, Model 2

	est	se	N	p	tau2	Q	Qp
outdegree (density)	-2.166	0.1496	20	0.000	0.199	18.300	0.502
reciprocity	0.609	0.0801	20	0.000	0.001	17.774	0.538
transitive triplets	-0.100	0.0272	20	0.000	0.057	23.570	0.213
indegree - popularity	0.118	0.0137	20	0.000	0.000	16.338	0.635
outdegree - activity	0.097	0.0092	20	0.000	0.017	22.067	0.281
indegree - activity	-0.018	0.0439	5	0.678	0.000	1.733	0.785
outdegree-trunc(1)	-2.555	1.1216	4	0.023	1.586	6.153	0.104
sex alter	0.026	0.0829	20	0.758	0.198	32.356	0.028
sex ego	0.115	0.0681	20	0.090	0.099	30.268	0.048
same sex	-0.514	0.0577	20	0.000	0.001	27.246	0.099
grade alter	0.005	0.0346	20	0.875	0.037	16.171	0.646
grade ego	0.040	0.0410	20	0.332	0.089	24.703	0.171
grade similarity	-0.004	0.1258	20	0.977	0.280	28.038	0.083
PGG alter	-0.037	0.0306	20	0.224	0.000	10.914	0.927
PGG ego	0.048	0.0320	20	0.135	0.071	24.542	0.176
PGG similarity	-0.096	0.1134	20	0.395	0.225	25.091	0.158
popularity alter	-0.132	0.0210	20	0.000	0.001	25.377	0.149
popularity ego	-0.025	0.0161	20	0.118	0.000	14.081	0.779
popularity ego x popularity alter	-0.016	0.0049	20	0.001	0.000	14.358	0.762
int. PGG ego x PGG similarity	0.238	0.0951	20	0.012	0.000	20.549	0.362

Table 30: Goodness of fit statistics, Negative and PGG, Model 2

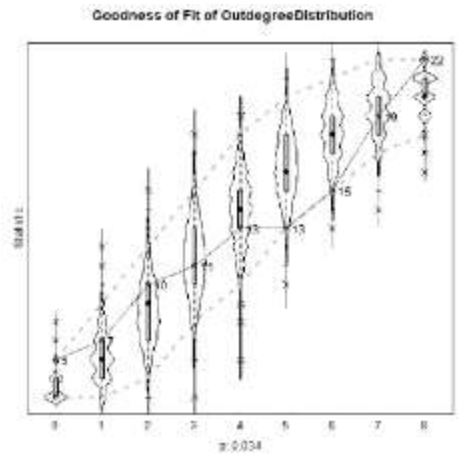
	Indegree distribution	Outdegree distribution	Geodesic Distance	Triad Census
1	0.727	0.415	0.935	0.717
2	1.000	0.813	0.962	0.848
3	0.661	0.970	1.000	0.672
4	0.278	0.355	0.186	0.096
5	0.303	0.146	0.056	0.967
6	0.971	0.034	0.015	0.997
7	0.016	0.745	0.187	0.759
8	0.950	0.518	0.072	0.982
9	0.898	0.454	0.091	0.884
10	0.898	0.009	0.804	0.807
11	0.655	0.804	0.901	0.694
12	0.209	0.479	0.866	0.997
13	0.988	0.069	0.042	0.519
14	0.384	0.232	0.657	0.928
15	0.983	0.099	0.361	0.278
16	0.688	0.746	0.072	0.792
17	0.989	0.883	0.990	1.000
18	0.046	0.327	0.235	0.583
19	0.993	0.818	0.532	0.820
20	0.502	0.685	0.935	0.578

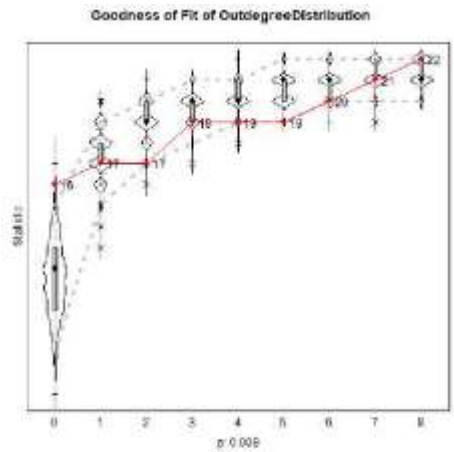




(a) Class 7

(b) Class 18





(c) Class 6

(d) Class 10

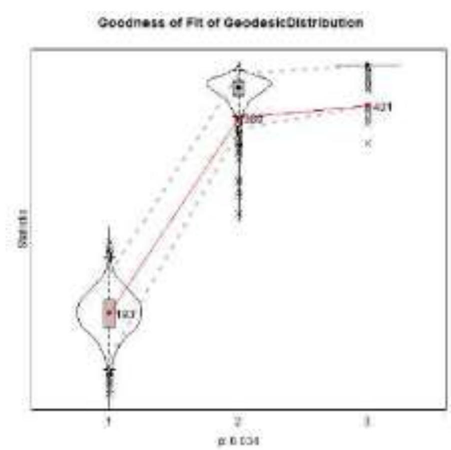


Figure 10: Classes with poor fit, Negative and PGG, Model 2

(e) Class 13

Table 31: Results of separate SAOMs for negative network and PGG, Model 2

	1	2	3	4	5	6	7	8	9	10
outdegree (density)	-1.968	2.376	0.246	-2.922	-2.376	-1.257	-2.122	-1.845	-1.535	-1.596
	[0.37]	[9.995]	[1.612]	[0.907]	[0.586]	[0.959]	[0.668]	[0.413]	[0.411]	[3.337]
reciprocity	0.718	2.533	0.005	0.458	1.135	0.623	1.172	0.502	0.301	0.288
	[0.223]	[1.641]	[0.492]	[0.54]	[0.301]	[0.416]	[0.38]	[0.219]	[0.278]	[1.006]
transitive triplets	-0.058	1.159	-0.163	0.382	-0.2	0.001	-0.099	-0.266	0.043	0.09
	[0.051]	[1.361]	[0.143]	[0.235]	[0.078]	[0.174]	[0.086]	[0.066]	[0.073]	[0.7]
indegree - popularity	0.094	-0.076	-0.308	0.07	0.085	0.12	0.094	0.123	0.032	-1.112
	[0.038]	[0.472]	[0.272]	[0.116]	[0.072]	[0.044]	[0.073]	[0.045]	[0.055]	[0.729]
outdegree - activity	0.073	-1.5	0.137	0.062	0.156	0.052	0.109	0.106	0.035	0.261
	[0.019]	[3.475]	[0.054]	[0.073]	[0.028]	[0.052]	[0.034]	[0.023]	[0.027]	[0.147]
indegree - activity	0	0	0	0	0	-0.164	0	0	0	-0.295
	[NA]	[NA]	[NA]	[NA]	[NA]	[0.134]	[NA]	[NA]	[NA]	[2.185]
outdegree-trunc(1)	0	0	0	1.33	0	0	0	-2.267	0	0
	[NA]	[NA]	[NA]	[1.999]	[NA]	[NA]	[NA]	[1.218]	[NA]	[NA]
sex alter	-0.012	-1.067	-1.959	0.111	-0.472	-0.161	0.125	0.371	0.239	-0.515
	[0.147]	[1.438]	[0.997]	[0.465]	[0.394]	[0.244]	[0.221]	[0.215]	[0.235]	[0.893]
sex ego	-0.028	3.689	-1.177	1.007	-0.328	-0.609	0.018	0.042	0.343	-0.488
	[0.156]	[10.643]	[0.533]	[0.542]	[0.365]	[0.484]	[0.266]	[0.167]	[0.24]	[0.971]
same sex	-0.532	-1.335	-0.654	0.077	-1.402	-0.154	-0.427	-0.675	-0.497	-1.175
	[0.158]	[1.113]	[0.432]	[0.393]	[0.398]	[0.204]	[0.208]	[0.167]	[0.204]	[1.126]
grade alter	0.089	-0.493	-0.059	0.074	0.14	0.077	0.27	-0.169	-0.27	0.098
	[0.133]	[0.688]	[0.237]	[0.224]	[0.173]	[0.103]	[0.19]	[0.094]	[0.149]	[0.49]
grade ego	0.205	-1.569	0.372	-0.217	0.222	0.096	0.553	-0.163	-0.199	-0.75
	[0.144]	[4.669]	[0.233]	[0.23]	[0.159]	[0.119]	[0.258]	[0.101]	[0.15]	[1.192]
grade similarity	-0.03	-0.323	0.519	-0.486	-0.503	-0.248	-0.21	-0.584	0.137	3.277
	[0.255]	[1.681]	[0.459]	[0.743]	[0.878]	[0.427]	[0.334]	[0.319]	[0.308]	[1.906]
PGG alter	0.03	0.7	-0.057	0.012	-0.094	0.087	-0.14	-0.169	-0.062	0.406
	[0.078]	[1.801]	[0.389]	[0.217]	[0.134]	[0.107]	[0.163]	[0.088]	[0.145]	[0.575]
PGG ego	-0.09	-6.041	0.621	0.155	-0.073	0.182	0.321	0.165	0	-0.199
	[0.082]	[15.39]	[0.309]	[0.157]	[0.134]	[0.141]	[0.225]	[0.093]	[0.092]	[0.5]
PGG similarity	-0.173	0.961	0.348	1.735	-0.686	0.749	0.272	-0.25	-0.445	1.991
	[0.293]	[1.758]	[0.819]	[0.856]	[0.464]	[0.407]	[0.519]	[0.352]	[0.314]	[2.393]
popularity alter	-0.098	-0.088	-0.808	-0.237	-0.269	-0.06	-0.33	-0.167	-0.14	-1.334
	[0.055]	[0.447]	[0.417]	[0.154]	[0.111]	[0.073]	[0.159]	[0.057]	[0.084]	[0.569]
popularity ego	0.017	-3.718	-0.28	0.073	-0.051	-0.245	-0.052	-0.037	-0.102	0.013
	[0.034]	[8.543]	[0.126]	[0.125]	[0.069]	[0.191]	[0.097]	[0.04]	[0.049]	[1.561]
popularity ego x popularity alter	-0.024	0.123	0.01	-0.11	0.017	-0.018	0.019	-0.014	0.003	0.046
	[0.008]	[0.379]	[0.036]	[0.052]	[0.035]	[0.019]	[0.038]	[0.015]	[0.017]	[0.102]
int. PGG ego x PGG similarity	0.105	-4.686	0.726	0.253	0.57	0	0.962	0.982	0.267	-1.354
	[0.278]	[9.916]	[0.63]	[0.799]	[0.485]	[0.366]	[0.582]	[0.325]	[0.365]	[1.295]
Overall maximum convergence ratio:	0.145	0.155	0.215	0.162	0.135	0.139	0.167	0.145	0.117	0.194

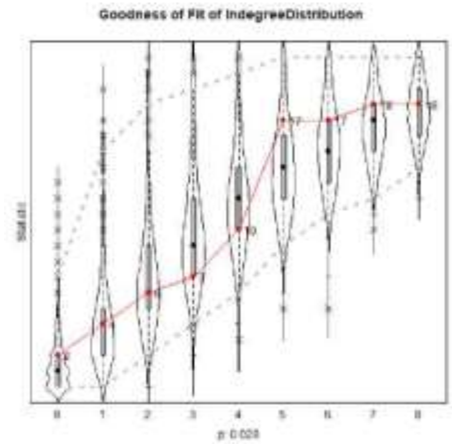
	11	12	13	14	15	16	17	18	19	20
outdegree (density)	-1.603	-2.87	-1.662	-2.989	-3.062	-2.55	-2.292	-2.322	-0.343	-1.973
	[0.694]	[0.556]	[0.833]	[0.497]	[0.498]	[0.982]	[0.894]	[0.399]	[1.333]	[0.889]
reciprocity	0.502	1.019	0.83	0.685	-0.086	0.436	0.442	0.665	-0.751	0.032
	[0.342]	[0.332]	[0.326]	[0.276]	[0.404]	[0.282]	[0.841]	[0.427]	[1.229]	[0.581]
transitive triplets	-0.213	-0.05	0.003	-0.117	-0.232	-0.116	-0.726	-0.024	-0.329	-0.298
	[0.092]	[0.07]	[0.114]	[0.06]	[0.184]	[0.089]	[0.974]	[0.086]	[0.505]	[0.244]
indegree - popularity	0.11	0.109	0.095	0.169	0.19	0.12	-0.103	0.168	-0.123	0.05
	[0.072]	[0.063]	[0.055]	[0.041]	[0.055]	[0.051]	[0.192]	[0.04]	[0.245]	[0.19]
outdegree - activity	0.088	0.1	0.065	0.109	0.17	0.101	0.032	0.078	0.159	0.22
	[0.041]	[0.027]	[0.037]	[0.028]	[0.045]	[0.028]	[0.079]	[0.032]	[0.115]	[0.075]
indegree - activity			-0.059			0.01		0.015	0	
	0 [NA]	0 [NA]	[0.105]	0 [NA]	0 [NA]	[0.101]	0 [NA]	[0.06]	[NA]	0 [NA]
outdegree-trunc(1)	-4.156						0		-4.418	
	[1.346]	0 [NA]	0 [NA]	0 [NA]	0 [NA]	0 [NA]	[NA]	0 [NA]	[1.929]	0 [NA]
sex alter	-0.347	-0.284	-0.215	0.265	1.241	0.035	-0.46	0.37	-1.387	-0.783
	[0.244]	[0.226]	[0.334]	[0.187]	[0.405]	[0.238]	[0.72]	[0.22]	[1.229]	[0.561]
sex ego	-0.121	0.101	0.101	0.235	0.802	0.377	1.749	0.419	-0.817	-0.304
	[0.235]	[0.228]	[0.285]	[0.191]	[0.312]	[0.254]	[0.907]	[0.211]	[0.688]	[0.44]
same sex	-0.31	-0.337	-0.446	-0.423	-1.021	-0.492	0.421	-0.976	-1.598	-1.54
	[0.23]	[0.217]	[0.275]	[0.198]	[0.332]	[0.2]	[0.561]	[0.29]	[0.782]	[0.551]
grade alter	0.009	0.027	0.107	0.01	0.043	0.046	0.007	-0.183	0.213	0.347
	[0.119]	[0.183]	[0.113]	[0.108]	[0.129]	[0.12]	[0.237]	[0.125]	[0.497]	[0.267]
grade ego	0.015	-0.075	-0.121	0.146	-0.032	0.223	-0.033	0.056	-0.094	
	[0.133]	[0.18]	[0.104]	[0.113]	[0.116]	[0.119]	[0.218]	[0.132]	[0.236]	0.2 [0.21]
grade similarity	-0.642	-0.664	-0.159	0.817	1.496	0.75	0.854	0.085	-1.438	0.183
	[0.436]	[0.406]	[0.41]	[0.443]	[0.704]	[0.406]	[0.998]	[0.394]	[1.125]	[0.737]
PGG alter	0.052	-0.167	0.022	-0.075	-0.154	-0.026	-0.241	-0.006	0.439	0.133
	[0.115]	[0.126]	[0.11]	[0.116]	[0.157]	[0.098]	[0.285]	[0.148]	[0.359]	[0.4]
PGG ego	0.153	-0.118	0.109	-0.145	0.017	0.198	-0.029	0.013	0.109	-0.065
	[0.081]	[0.118]	[0.099]	[0.09]	[0.095]	[0.087]	[0.22]	[0.059]	[0.26]	[0.265]
PGG similarity	-0.537	-0.467	0.072	0.531	-0.578	-0.188	-0.316	-0.339	0.047	1.614
	[0.328]	[0.444]	[0.403]	[0.318]	[0.462]	[0.319]	[0.851]	[0.253]	[1.412]	[1.082]
popularity alter	-0.269	-0.118	-0.102	-0.048	-0.06	-0.222	-0.272	-0.003	-0.762	-0.498
	[0.115]	[0.059]	[0.097]	[0.066]	[0.12]	[0.087]	[0.18]	[0.075]	[0.41]	[0.237]
popularity ego	-0.065	0.002	0.068	0	-0.019	0.099	-0.086	0.092	0.016	-0.127
	[0.072]	[0.044]	[0.146]	[0.06]	[0.086]	[0.211]	[0.117]	[0.109]	[0.123]	[0.156]
popularity ego x popularity alter	-0.023	-0.004	-0.019	-0.038	-0.024	0.024	-0.031	-0.027	-0.024	-0.101
	[0.021]	[0.014]	[0.031]	[0.023]	[0.038]	[0.028]	[0.072]	[0.036]	[0.044]	[0.058]
int. PGG ego x PGG similarity	0.389	0.431	0.013	0.112	0.426	0.247	-0.748	-0.357	-0.384	-1.892
	[0.301]	[0.382]	[0.428]	[0.32]	[0.477]	[0.286]	[0.719]	[0.367]	[0.881]	[1.043]
Overall maximum convergence ratio:	0.155	0.166	0.148	0.094	0.165	0.170	0.137	0.143	0.157	0.165

Table 32: Results of the meta-analysis for negative network and DG, Model 1

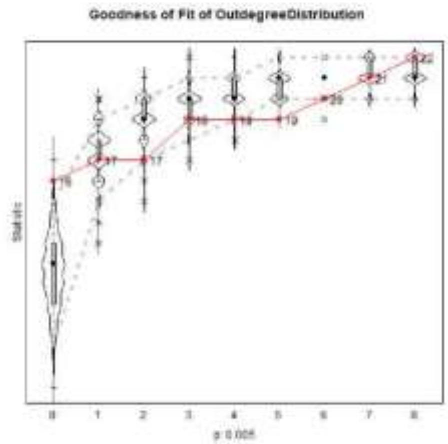
	est	se	N	p	tau2	Q	Qp
outdegree (density)	-2.165	0.1536	20	0.000	0.2434	19.936	0.398
reciprocity	0.594	0.0783	20	0.000	0.0000	16.767	0.606
transitive triplets	-0.106	0.0262	20	0.000	0.0483	21.924	0.288
indegree - popularity	0.120	0.0136	20	0.000	0.0016	20.204	0.382
outdegree - activity	0.099	0.0094	20	0.000	0.0184	24.713	0.170
indegree - activity	-0.011	0.0453	5	0.814	0.0000	1.627	0.804
outdegree-trunc(1)	-2.887	1.0315	4	0.005	1.2884	5.508	0.138
sex alter	0.025	0.0729	20	0.737	0.1311	34.714	0.015
sex ego	0.166	0.0625	20	0.008	0.0000	23.010	0.237
same sex	-0.505	0.0587	20	0.000	0.0004	25.987	0.131
grade alter	0.015	0.0370	20	0.682	0.0628	20.632	0.357
grade ego	-0.004	0.0359	20	0.905	0.0464	16.746	0.607
grade similarity	0.014	0.1300	20	0.913	0.3011	28.379	0.076
DG alter	-0.007	0.0167	20	0.688	0.0251	21.209	0.325
DG ego	0.002	0.0210	20	0.937	0.0467	26.018	0.130
DG similarity	-0.201	0.0829	20	0.015	0.0161	26.436	0.118
popularity alter	-0.138	0.0231	20	0.000	0.0324	25.162	0.155
popularity ego	-0.003	0.0151	20	0.852	0.0000	11.102	0.920
popularity ego x popularity alter	-0.017	0.0049	20	0.000	0.0000	12.882	0.845

Table 33: Goodness of fit statistics, Negative and DG, Model 1

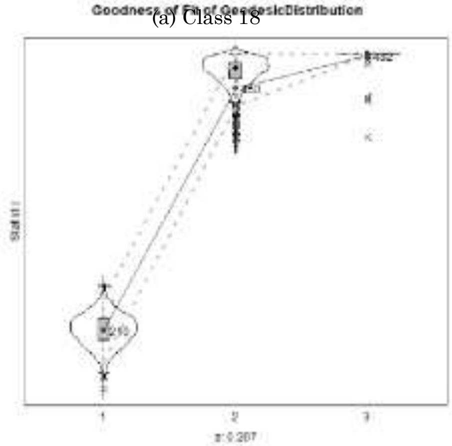
	Indegree distribution	Outdegree distribution	Geodesic Distance	Triad Census
1	0.764	0.374	0.937	0.694
2	0.912	0.944	0.981	0.876
3	0.649	0.935	1.000	0.729
4	0.349	0.324	0.258	0.130
5	0.316	0.160	0.034	0.978
6	0.949	0.052	0.014	0.998
7	0.060	0.818	0.612	0.928
8	0.924	0.518	0.062	0.982
9	0.878	0.474	0.098	0.893
10	0.904	0.006	0.764	0.845
11	0.533	0.791	0.900	0.791
12	0.222	0.488	0.877	0.992
13	0.994	0.226	0.040	0.618
14	0.413	0.202	0.402	0.940
15	0.983	0.186	0.334	0.286
16	0.738	0.798	0.048	0.753
17	0.989	0.863	0.918	1.000
18	0.021	0.387	0.478	0.498
19	0.986	0.760	0.230	0.619
20	0.475	0.613	0.949	0.708



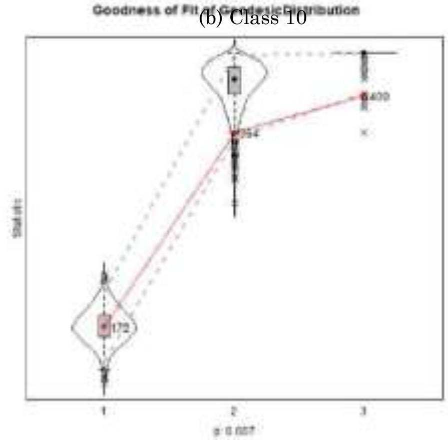
(a) Class 18



(b) Class 10



(c) Class 5



(d) Class 6

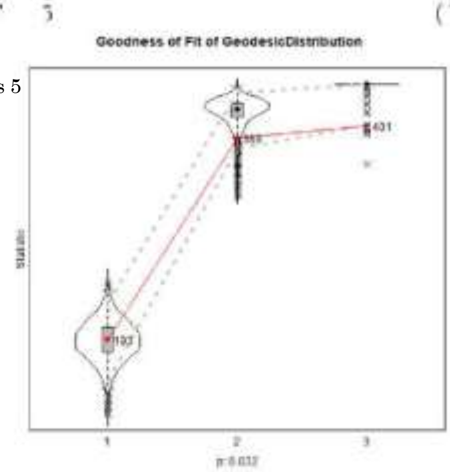


Figure 11: Classes with poor fit, Negative and DG, Model 1

(e) Class 13

Table 34: Results of separate SAOMs for negative network and DG, Model 1

	1	2	3	4	5	6	7	8	9	10
outdegree (density)	-2.043	5.324	-0.565	-2.833	-2.295	-0.831	-2.529	-1.611	-1.535	-1.443
	[0.345]	[7.555]	[1.431]	[0.9]	[0.595]	[1.513]	[0.673]	[0.397]	[0.423]	[4.636]
reciprocity	0.68	1.779	0.101	0.54	1.102	0.65	1.195	0.562	0.336	0.112
	[0.215]	[1.116]	[0.459]	[0.486]	[0.301]	[0.407]	[0.421]	[0.213]	[0.259]	[1.103]
transitive triplets	-0.063	1.072	-0.193	0.328	-0.201	0.033	-0.156	-0.233	0.037	0.119
	[0.05]	[1.156]	[0.165]	[0.253]	[0.079]	[0.169]	[0.115]	[0.065]	[0.073]	[0.53]
indegree - popularity	0.103	-0.949	-0.304	0.075	0.083	0.129	0.122	0.075	0.024	-1.226
	[0.037]	[0.689]	[0.26]	[0.118]	[0.07]	[0.038]	[0.082]	[0.048]	[0.056]	[0.891]
outdegree - activity	0.076	-1.943	0.167	0.07	0.152	0.028	0.158	0.095	0.036	0.252
	[0.017]	[2.461]	[0.053]	[0.075]	[0.028]	[0.071]	[0.052]	[0.021]	[0.026]	[0.144]
indegree - activity	0	0	0	0	0	-0.229	0	0	0	-0.148
	[NA]	[NA]	[NA]	[NA]	[NA]	[0.218]	[NA]	[NA]	[NA]	[3.104]
outdegree-trunc(1)	0	0	0	1.211	0	0	0	-2.321	0	0
	[NA]	[NA]	[NA]	[2.291]	[NA]	[NA]	[NA]	[1.222]	[NA]	[NA]
sex alter	-0.03	-3.259	-1.894	0.007	-0.428	-0.07	-0.024	0.227	0.225	-0.441
	[0.153]	[1.873]	[0.807]	[0.426]	[0.391]	[0.263]	[0.291]	[0.195]	[0.22]	[1.032]
sex ego	-0.013	-0.543	-0.237	0.887	-0.372	-0.528	-0.264	0.163	0.36	-0.44
	[0.154]	[2.082]	[0.382]	[0.503]	[0.365]	[0.622]	[0.333]	[0.162]	[0.236]	[0.93]
same sex	-0.541	-1.426	-0.481	0.081	-1.398	-0.179	-0.501	-0.659	-0.51	-1.305
	[0.157]	[0.988]	[0.485]	[0.389]	[0.422]	[0.204]	[0.26]	[0.166]	[0.21]	[0.787]
grade alter	0.114	-0.445	-0.006	0.157	0.185	0.053	0.363	-0.265	-0.266	0.132
	[0.137]	[0.478]	[0.233]	[0.213]	[0.18]	[0.097]	[0.206]	[0.108]	[0.145]	[0.53]
grade ego	0.141	1.185	-0.075	-0.197	0.217	0.105	0.78	-0.101	-0.198	-0.708
	[0.14]	[1.388]	[0.189]	[0.212]	[0.16]	[0.116]	[0.688]	[0.096]	[0.153]	[1.327]
grade similarity	-0.011	-0.284	0.674	-0.348	-0.232	-0.198	-0.176	-0.661	0.174	3.178
	[0.259]	[1.377]	[0.465]	[0.725]	[0.916]	[0.446]	[0.371]	[0.289]	[0.29]	[2.011]
DG alter	0.029	-0.75	0.087	0.037	-0.137	-0.007	-0.079	-0.104	0.019	0.211
	[0.044]	[0.415]	[0.139]	[0.136]	[0.115]	[0.047]	[0.099]	[0.057]	[0.055]	[0.25]
DG ego	-0.068	-1.394	-0.217	0.088	-0.144	0.056	0.439	0.076	-0.084	0.016
	[0.046]	[1.76]	[0.13]	[0.119]	[0.13]	[0.083]	[0.391]	[0.045]	[0.059]	[0.337]
DG similarity	-0.144	-1.27	-0.428	0.14	0.122	-0.472	1.78	-0.054	0.012	2.248
	[0.268]	[1.508]	[0.51]	[0.721]	[0.557]	[0.29]	[1.051]	[0.251]	[0.321]	[1.495]
popularity alter	-0.1	-0.272	-0.794	-0.252	-0.266	-0.05	-0.26	-0.202	-0.157	-1.43
	[0.055]	[0.438]	[0.363]	[0.158]	[0.1]	[0.073]	[0.13]	[0.065]	[0.091]	[0.675]
popularity ego	0.027	-2.26	-0.106	0.07	-0.03	-0.331	-0.182	-0.024	-0.081	0.117
	[0.03]	[2.948]	[0.096]	[0.117]	[0.068]	[0.322]	[0.303]	[0.039]	[0.05]	[1.99]
popularity ego x popularity alter	-0.026	0.192	0.004	-0.108	0.018	-0.017	-0.023	-0.017	0.003	0.071
	[0.009]	[0.346]	[0.035]	[0.049]	[0.035]	[0.02]	[0.035]	[0.016]	[0.016]	[0.11]
Overall maximum convergence ratio:	0.115	0.130	0.161	0.143	0.125	0.174	0.112	0.185	0.105	0.133

CH

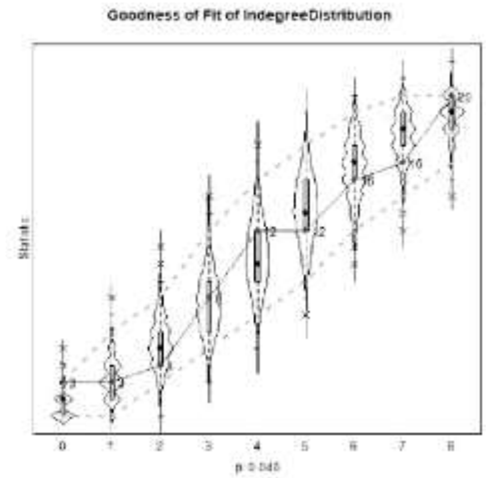
	11	12	13	14	15	16	17	18	19	20
outdegree (density)	-1.994	-3.041	-1.596	-2.686	-3.2	-2.564	-2.475	-2.362	-0.497	-1.647
	[0.585]	[0.552]	[0.976]	[0.447]	[0.547]	[1.078]	[0.877]	[0.411]	[0.959]	[0.895]
reciprocity	0.509	0.982	0.788	0.629	-0.06	0.379	0.059	0.643	-1.194	-0.021
	[0.331]	[0.335]	[0.306]	[0.269]	[0.378]	[0.322]	[0.838]	[0.389]	[1.613]	[0.539]
transitive triplets	-0.233	-0.039	-0.003	-0.112	-0.254	-0.119	-0.664	-0.048	-0.678	-0.313
	[0.093]	[0.068]	[0.128]	[0.061]	[0.184]	[0.082]	[0.762]	[0.1]	[0.59]	[0.235]
indegree - popularity	0.157	0.109	0.095	0.165	0.203	0.117	-0.062	0.182	-0.055	0
	[0.061]	[0.063]	[0.059]	[0.042]	[0.057]	[0.048]	[0.176]	[0.041]	[0.199]	[0.207]
outdegree - activity	0.097	0.102	0.069	0.093	0.177	0.11	0.035	0.082	0.203	0.221
	[0.038]	[0.025]	[0.042]	[0.025]	[0.051]	[0.027]	[0.076]	[0.034]	[0.138]	[0.067]
indegree - activity	0	0	-0.08	0	0	0.003	0	0.017	0	0
	[NA]	[NA]	[0.116]	[NA]	[NA]	[0.125]	[NA]	[0.055]	[NA]	[NA]
outdegree-trunc(1)	-4.534	0	0	0	0	0	0	0	-4.468	0
	[1.457]	[NA]	[NA]	[NA]	[NA]	[NA]	[NA]	[NA]	[1.739]	[NA]
sex alter	-0.119	-0.347	-0.18	0.241	1.229	0.065	-1.327	0.413	-2.241	-0.561
	[0.217]	[0.26]	[0.28]	[0.196]	[0.394]	[0.23]	[0.784]	[0.235]	[1.743]	[0.519]
sex ego	0.091	0.269	0.148	0.216	0.786	0.387	1.826	0.497	-1.029	-0.051
	[0.234]	[0.256]	[0.249]	[0.187]	[0.314]	[0.29]	[0.923]	[0.242]	[0.857]	[0.384]
same sex	-0.275	-0.242	-0.456	-0.419	-1.029	-0.493	0.443	-0.997	-1.392	-1.558
O	[0.218]	[0.22]	[0.287]	[0.189]	[0.352]	[0.194]	[0.65]	[0.319]	[0.793]	[0.599]
grade alter	0.039	0.073	0.127	-0.026	0.038	0.037	-0.002	-0.085	0.59	0.245
	[0.113]	[0.202]	[0.115]	[0.106]	[0.129]	[0.112]	[0.248]	[0.112]	[0.502]	[0.236]
grade ego	0.007	-0.106	-0.146	0.004	-0.043	0.221	-0.094	-0.024	-0.031	0.087
	[0.126]	[0.187]	[0.097]	[0.119]	[0.121]	[0.111]	[0.246]	[0.146]	[0.25]	[0.214]
grade similarity	-0.494	-0.727	-0.223	0.922	1.385	0.586	1.348	0.181	-1.377	0.015
	[0.43]	[0.436]	[0.45]	[0.494]	[0.707]	[0.388]	[1.055]	[0.397]	[1.072]	[0.698]
DG alter	-0.021	-0.002	0.069	0.044	0.071	0.036	-0.142	-0.092	0.08	0.19
	[0.055]	[0.058]	[0.06]	[0.053]	[0.109]	[0.048]	[0.111]	[0.042]	[0.172]	[0.175]
DG ego	-0.041	-0.085	0.106	0.095	0.224	-0.038	0.027	-0.015	0.068	0.073
	[0.058]	[0.067]	[0.076]	[0.058]	[0.115]	[0.044]	[0.123]	[0.048]	[0.105]	[0.165]
DG similarity	-0.003	-0.667	-0.21	-0.012	-1.079	-0.336	-1.762	0.252	-0.703	-0.107
	[0.297]	[0.339]	[0.301]	[0.3]	[0.638]	[0.252]	[0.538]	[0.27]	[0.48]	[0.749]
popularity alter	-0.221	-0.092	-0.125	-0.043	-0.054	-0.246	-0.189	-0.015	-0.847	-0.391
	[0.1]	[0.061]	[0.1]	[0.061]	[0.124]	[0.091]	[0.164]	[0.078]	[0.462]	[0.194]
popularity ego	-0.059	0.031	0.055	0.029	0.022	0.046	-0.051	0.126	0.015	-0.049
	[0.065]	[0.04]	[0.158]	[0.05]	[0.084]	[0.234]	[0.122]	[0.12]	[0.123]	[0.093]
popularity ego x popularity alter	-0.019	-0.011	-0.018	-0.027	-0.017	0.028	-0.059	-0.026	-0.02	-0.058
	[0.02]	[0.014]	[0.029]	[0.022]	[0.037]	[0.027]	[0.071]	[0.036]	[0.04]	[0.04]
Overall maximum convergence ratio:	0.153	0.102	0.148	0.115	0.126	0.139	0.167	0.203	0.216	0.173

Table 35: Results of the meta-analysis for negative network and DG, Model 2

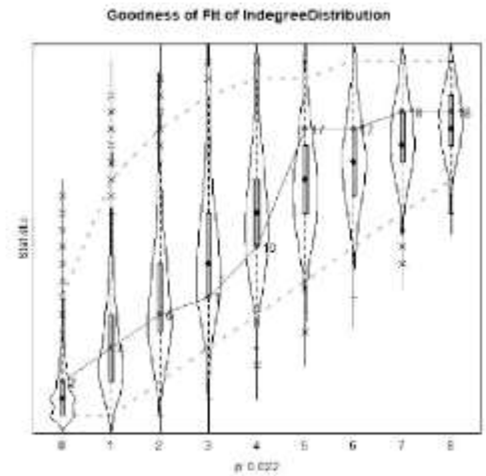
	est	se	N	p	tau2	Q	Qp
outdegree (density)	-2.175	0.1567	20	0.000	0.244	19.873	0.402
reciprocity	0.584	0.0814	20	0.000	0.001	17.516	0.555
transitive triplets	-0.105	0.0277	20	0.000	0.055	23.515	0.215
indegree - popularity	0.118	0.0140	20	0.000	0.000	18.314	0.502
outdegree - activity	0.098	0.0095	20	0.000	0.019	24.985	0.161
indegree - activity	-0.014	0.0460	5	0.761	0.000	1.151	0.886
outdegree-trunc(1)	-2.572	1.2347	4	0.037	1.780	6.238	0.101
sex alter	0.028	0.0693	20	0.686	0.100	33.681	0.020
sex ego	0.174	0.0628	20	0.006	0.002	24.712	0.170
same sex	-0.500	0.0593	20	0.000	0.044	28.332	0.077
grade alter	0.016	0.0344	20	0.636	0.029	18.877	0.465
grade ego	0.003	0.0385	20	0.944	0.065	19.327	0.436
grade similarity	0.019	0.1367	20	0.888	0.342	31.224	0.038
DG alter	-0.007	0.0252	20	0.768	0.000	13.061	0.835
DG ego	-0.002	0.0188	20	0.922	0.028	20.969	0.339
DG similarity	-0.204	0.0862	20	0.018	0.001	28.784	0.069
popularity alter	-0.143	0.0238	20	0.000	0.035	25.835	0.135
popularity ego	-0.004	0.0157	20	0.794	0.000	11.074	0.921
popularity ego x popularity alter	-0.017	0.0048	20	0.000	0.001	13.823	0.794
int. DG ego x DG similarity	0.018	0.0833	20	0.828	0.173	27.510	0.093

Table 36: Goodness of fit statistics, Negative and DG, Model 2

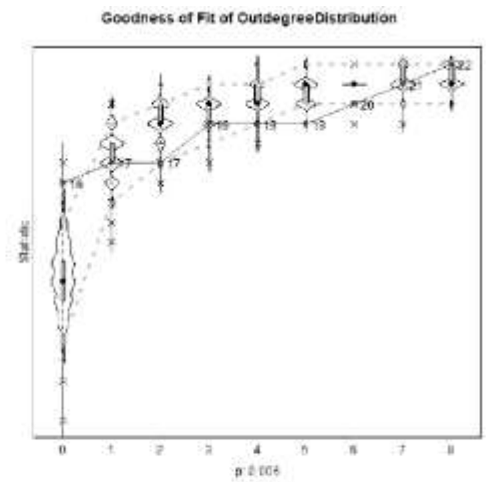
	Indegree distribution	Outdegree distribution	Geodesic Distance	Triad Census
1	0.756	0.352	0.925	0.748
2	0.825	0.949	0.985	0.972
3	0.632	0.950	0.990	0.717
4	0.300	0.333	0.277	0.122
5	0.367	0.288	0.105	0.982
6	0.970	0.054	0.024	0.997
7	0.048	0.814	0.484	0.814
8	0.913	0.564	0.078	0.970
9	0.877	0.514	0.086	0.868
10	0.909	0.007	0.772	0.843
11	0.559	0.807	0.829	0.789
12	0.239	0.487	0.892	0.991
13	0.997	0.226	0.058	0.656
14	0.413	0.196	0.425	0.942
15	0.978	0.096	0.284	0.228
16	0.781	0.741	0.076	0.745
17	0.986	0.890	0.938	1.000
18	0.032	0.383	0.383	0.469
19	0.982	0.781	0.286	0.630
20	0.540	0.568	0.942	0.683



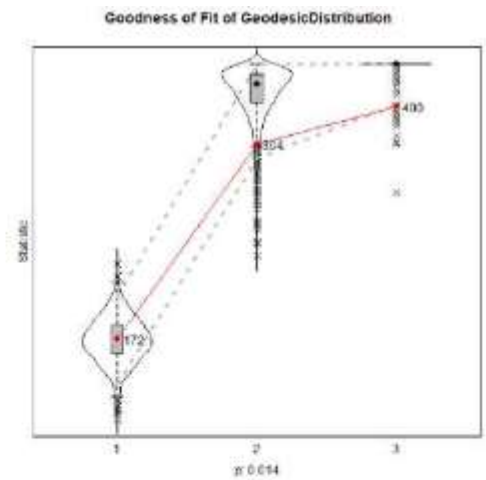
(a) Class 7



(b) Class 18



(c) Class 10



(d) Class 6

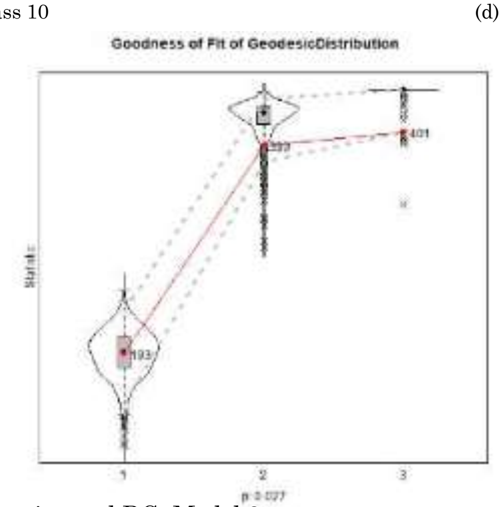


Figure 12: Classes with poor fit, Negative and DG, Model 2

(e) Class 13

Table 37: Results of separate SAOMs for negative network and DG, Model 2

	1	2	3	4	5	6	7	8	9	10
outdegree (density)	-2.082	6.667	-0.615	-2.738	-2.663	-1.021	-2.691	-1.622	-1.537	-1.784
	[0.366]	[8.002]	[1.549]	[0.838]	[0.656]	[1.264]	[0.788]	[0.387]	[0.412]	[3.34]
reciprocity	0.69	2.675	0.12	0.532	1.096	0.638	1.254	0.576	0.332	-0.04
	[0.229]	[2.065]	[0.467]	[0.517]	[0.311]	[0.445]	[0.495]	[0.21]	[0.273]	[1.213]
transitive triplets	-0.064	1.753	-0.205	0.35	-0.218	0.013	-0.169	-0.232	0.036	0.115
	[0.055]	[1.943]	[0.142]	[0.213]	[0.085]	[0.247]	[0.14]	[0.065]	[0.066]	[0.708]
indegree - popularity	0.105	-1.401	-0.302	0.064	0.086	0.13	0.132	0.078	0.026	-1.189
	[0.039]	[1.162]	[0.248]	[0.12]	[0.069]	[0.047]	[0.09]	[0.047]	[0.054]	[1.091]
outdegree - activity	0.077	-2.366	0.171	0.063	0.156	0.037	0.165	0.096	0.036	0.265
	[0.019]	[2.748]	[0.049]	[0.069]	[0.03]	[0.081]	[0.069]	[0.021]	[0.025]	[0.157]
indegree - activity	0	0	0	0	0	-0.202	0	0	0	-0.139
	[NA]	[NA]	[NA]	[NA]	[NA]	[0.232]	[NA]	[NA]	[NA]	[2.024]
outdegree-trunc(1)	0	0	0	1.299	0	0	0	-2.374	0	0
	[NA]	[NA]	[NA]	[1.957]	[NA]	[NA]	[NA]	[1.292]	[NA]	[NA]
sex alter	-0.024	-5.768	-1.908	0.003	-0.384	-0.078	-0.03	0.232	0.233	-0.502
	[0.156]	[3.966]	[0.764]	[0.42]	[0.421]	[0.283]	[0.283]	[0.191]	[0.219]	[1.003]
sex ego	-0.014	-1.786	-0.222	0.898	-0.463	-0.482	-0.15	0.116	0.372	-0.408
	[0.153]	[2.605]	[0.384]	[0.457]	[0.416]	[0.574]	[0.321]	[0.157]	[0.229]	[0.818]
same sex	-0.55	-2.691	-0.447	0.088	-1.404	-0.172	-0.522	-0.619	-0.506	-1.554
O	[0.154]	[2.186]	[0.483]	[0.379]	[0.447]	[0.206]	[0.27]	[0.166]	[0.203]	[1.106]
grade alter	0.122	-0.715	-0.003	0.14	0.156	0.059	0.404	-0.265	-0.267	0.057
	[0.136]	[0.753]	[0.222]	[0.209]	[0.168]	[0.095]	[0.27]	[0.119]	[0.149]	[0.58]
grade ego	0.166	1.633	-0.074	-0.227	0.339	0.091	0.849	-0.113	-0.196	-0.601
	[0.147]	[1.84]	[0.174]	[0.209]	[0.181]	[0.11]	[0.854]	[0.102]	[0.147]	[1.291]
grade similarity	-0.036	-0.362	0.627	-0.292	0.345	-0.193	-0.193	-0.687	0.163	3.607
	[0.257]	[1.802]	[0.502]	[0.741]	[0.927]	[0.414]	[0.391]	[0.298]	[0.295]	[1.995]
DG alter	-0.018	-0.071	0.046	0.108	-0.502	0.05	0.116	0.008	0.042	0.518
	[0.068]	[0.639]	[0.143]	[0.185]	[0.241]	[0.092]	[0.219]	[0.077]	[0.075]	[0.453]
DG ego	-0.056	-2.033	-0.201	0.053	0.023	0.046	0.388	0.045	-0.092	0.02
	[0.048]	[2.014]	[0.122]	[0.134]	[0.161]	[0.079]	[0.436]	[0.044]	[0.064]	[0.306]
DG similarity	-0.083	-3.992	-0.362	0.066	1.159	-0.407	1.678	-0.084	-0.022	2.381
	[0.263]	[4.288]	[0.517]	[0.71]	[0.814]	[0.295]	[1.261]	[0.265]	[0.329]	[1.275]
popularity alter	-0.1	-0.664	-0.806	-0.257	-0.269	-0.051	-0.27	-0.192	-0.155	-1.501
	[0.058]	[0.741]	[0.322]	[0.152]	[0.097]	[0.089]	[0.134]	[0.065]	[0.089]	[0.782]
popularity ego	0.026	-2.601	-0.1	0.081	-0.055	-0.295	-0.187	-0.02	-0.082	0.108
	[0.031]	[3.686]	[0.094]	[0.117]	[0.072]	[0.317]	[0.312]	[0.04]	[0.052]	[1.463]
popularity ego x popularity alter	-0.025	0.479	0.005	-0.108	0.011	-0.015	-0.03	-0.014	0.004	0.064
	[0.008]	[0.475]	[0.032]	[0.05]	[0.036]	[0.02]	[0.049]	[0.015]	[0.017]	[0.103]
int. DG ego x DG similarity	0.188	-3.442	0.239	-0.365	1.094	-0.178	-0.887	-0.41	-0.102	-1.403
	[0.206]	[2.454]	[0.326]	[0.559]	[0.621]	[0.262]	[0.73]	[0.206]	[0.223]	[1.117]
Overall maximum convergence ratio:	0.118	0.184	0.158	0.134	0.155	0.105	0.162	0.161	0.151	0.179

g

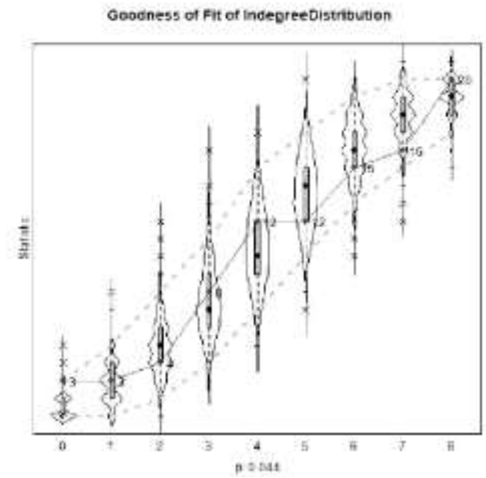
	11	12	13	14	15	16	17	18	19	20
outdegree (density)	-1.885	-3.072	-1.606	-2.736	-3.263	-2.498	-2.51	-2.31	-0.482	-1.817
	[0.586]	[0.587]	[0.915]	[0.467]	[0.604]	[0.957]	[0.752]	[0.458]	[1.044]	[0.878]
reciprocity	0.488	1.009	0.81	0.634	-0.065	0.395	0.072	0.697	-1.178	-0.07
	[0.324]	[0.35]	[0.346]	[0.286]	[0.386]	[0.324]	[0.862]	[0.376]	[1.093]	[0.557]
transitive triplets	-0.234	-0.039	-0.006	-0.116	-0.259	-0.111	-0.68	-0.04	-0.654	-0.306
	[0.107]	[0.071]	[0.12]	[0.06]	[0.198]	[0.075]	[0.822]	[0.096]	[0.563]	[0.231]
indegree - popularity	0.154	0.112	0.094	0.166	0.201	0.116	-0.066	0.18	-0.053	0.019
	[0.065]	[0.064]	[0.06]	[0.04]	[0.06]	[0.045]	[0.173]	[0.044]	[0.196]	[0.195]
outdegree - activity	0.095	0.103	0.07	0.095	0.179	0.109	0.035	0.08	0.201	0.223
	[0.04]	[0.027]	[0.037]	[0.025]	[0.052]	[0.026]	[0.062]	[0.034]	[0.118]	[0.071]
indegree - activity			-0.085			-0.007		0.01	0	
	0 [NA]	0 [NA]	[0.126]	0 [NA]	0 [NA]	[0.106]	0 [NA]	[0.057]	[NA]	0 [NA]
outdegree-trunc(1)	-4.57						0		-4.436	
	[1.652]	0 [NA]	0 [NA]	0 [NA]	0 [NA]	0 [NA]	[NA]	0 [NA]	[2.032]	0 [NA]
sex alter	-0.131	-0.36	-0.14	0.238	1.226	0.066	-1.337	0.412	-2.177	-0.512
	[0.223]	[0.25]	[0.308]	[0.2]	[0.427]	[0.212]	[0.738]	[0.238]	[1.464]	[0.526]
sex ego	0.045	0.306	0.237	0.227	0.781	0.366	1.944	0.52	-1.017	0.052
	[0.231]	[0.27]	[0.26]	[0.194]	[0.326]	[0.288]	[0.813]	[0.252]	[0.916]	[0.427]
same sex	-0.259	-0.208	-0.45	-0.404	-1.029	-0.499	0.457	-0.997	-1.359	-1.577
	[0.207]	[0.226]	[0.263]	[0.199]	[0.324]	[0.19]	[0.597]	[0.282]	[0.725]	[0.561]
grade alter	0.04	0.074	0.139	-0.026	0.038	0.039	-0.028	-0.086	0.575	0.288
	[0.118]	[0.19]	[0.119]	[0.108]	[0.121]	[0.115]	[0.23]	[0.117]	[0.501]	[0.237]
grade ego		-0.115	-0.141	0.015	-0.03	0.225	-0.056	-0.015	-0.023	0.19
	0 [0.135]	[0.187]	[0.098]	[0.123]	[0.121]	[0.109]	[0.253]	[0.156]	[0.272]	[0.236]
grade similarity	-0.508	-0.774	-0.171	0.919	1.406	0.587	1.498	0.218	-1.366	-0.141
	[0.423]	[0.404]	[0.451]	[0.475]	[0.694]	[0.368]	[1.069]	[0.409]	[0.953]	[0.752]
DG alter	-0.115	-0.109	-0.12	-0.052	0.016	0.08	-0.007	-0.03	0.203	0.065
	[0.095]	[0.1]	[0.117]	[0.094]	[0.152]	[0.086]	[0.157]	[0.098]	[0.203]	[0.219]
DG ego	-0.02	-0.077	0.132	0.103	0.256	-0.043	-0.035	-0.019	0.046	0.15
	[0.062]	[0.069]	[0.084]	[0.063]	[0.143]	[0.044]	[0.132]	[0.049]	[0.107]	[0.217]
DG similarity	-0.145	-0.8	-0.234	-0.017	-0.98	-0.345	-2.112	0.19	-0.965	-0.231
	[0.313]	[0.345]	[0.317]	[0.291]	[0.66]	[0.271]	[0.65]	[0.269]	[0.612]	[0.848]
popularity alter	-0.228	-0.086	-0.134	-0.041	-0.06	-0.241	-0.196	-0.02	-0.824	-0.374
	[0.102]	[0.059]	[0.106]	[0.062]	[0.131]	[0.086]	[0.164]	[0.076]	[0.434]	[0.184]
popularity ego	-0.066	0.037	0.049	0.037	0.024	0.03	-0.076	0.117	-0.011	-0.05
	[0.067]	[0.043]	[0.163]	[0.053]	[0.084]	[0.194]	[0.133]	[0.137]	[0.128]	[0.099]
popularity ego x popularity alter	-0.019	-0.011	-0.019	-0.026	-0.016	0.028	-0.06	-0.026	-0.025	-0.065
	[0.02]	[0.014]	[0.032]	[0.023]	[0.04]	[0.027]	[0.069]	[0.037]	[0.041]	[0.043]
int. DG ego x DG similarity	0.292	0.366	0.635	0.35	0.213	-0.146	-0.528	-0.172	-0.395	0.611
	[0.224]	[0.268]	[0.349]	[0.271]	[0.42]	[0.228]	[0.408]	[0.254]	[0.407]	[0.708]
Overall maximum convergence ratio:	0.118	0.138	0.171	0.097	0.148	0.114	0.127	0.140	0.177	0.124

Table 38: Results of the meta-analysis for negative network and TG, Model 1

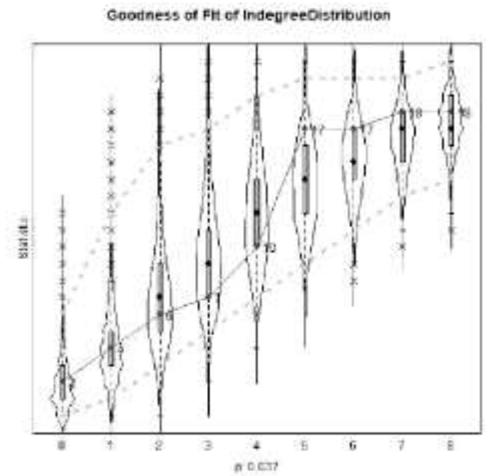
	est	se	N	p	tau2	Q	Qp
outdegree (density)	-2.098	0.1547	20	0.000	0.252	20.728	0.352
reciprocity	0.588	0.0816	20	0.000	0.000	16.830	0.601
transitive triplets	-0.104	0.0266	20	0.000	0.052	22.547	0.258
indegree - popularity	0.110	0.0142	20	0.000	0.001	17.220	0.575
outdegree - activity	0.098	0.0097	20	0.000	0.021	25.470	0.146
indegree - activity	-0.008	0.0465	4	0.863	0.000	2.165	0.539
outdegree-trunc(1)	-2.808	1.1209	4	0.012	1.548	6.060	0.109
sex alter	0.020	0.0830	20	0.810	0.195	36.574	0.009
sex ego	0.156	0.0771	20	0.043	0.163	32.600	0.027
same sex	-0.507	0.0630	20	0.000	0.091	29.936	0.053
grade alter	0.013	0.0341	20	0.692	0.000	16.551	0.620
grade ego	0.000	0.0331	20	0.990	0.000	13.231	0.827
grade similarity	0.003	0.1244	20	0.980	0.263	26.510	0.117
TG alter	-0.008	0.0207	20	0.705	0.054	34.661	0.015
TG ego	-0.011	0.0240	20	0.638	0.069	35.745	0.011
TG similarity	-0.162	0.1073	20	0.131	0.002	21.701	0.299
popularity alter	-0.146	0.0219	20	0.000	0.001	24.474	0.179
popularity ego	-0.016	0.0161	20	0.321	0.000	10.177	0.948
popularity ego x popularity alter	-0.014	0.0050	20	0.004	0.000	14.168	0.774

Table 39: Goodness of fit statistics, Negative and TG, Model 1

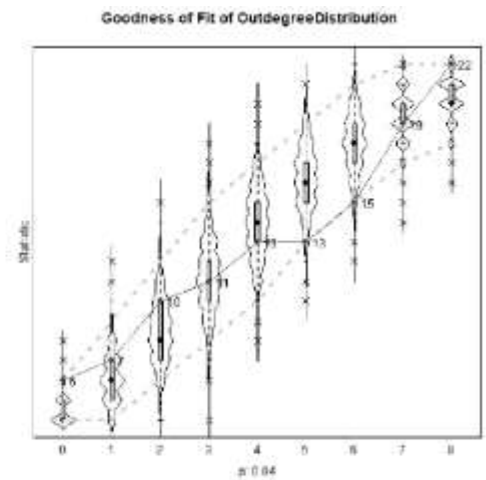
	Indegree distribution	Outdegree distribution	Geodesic Distance	Triad Census
1	0.773	0.560	0.962	0.647
2	0.936	0.881	0.994	0.779
3	0.643	0.961	1.000	0.758
4	0.172	0.282	0.108	0.125
5	0.325	0.112	0.080	0.981
6	0.970	0.040	0.023	0.996
7	0.044	0.791	0.291	0.877
8	0.947	0.519	0.042	0.979
9	0.895	0.449	0.077	0.913
10	0.403	0.017	0.604	0.833
11	0.683	0.722	0.894	0.725
12	0.213	0.502	0.922	0.995
13	0.992	0.133	0.043	0.544
14	0.463	0.147	0.286	0.919
15	0.989	0.275	0.488	0.337
16	0.764	0.851	0.073	0.729
17	0.988	0.845	0.961	1.000
18	0.037	0.314	0.577	0.447
19	0.982	0.741	0.282	0.653
20	0.472	0.636	0.938	0.800



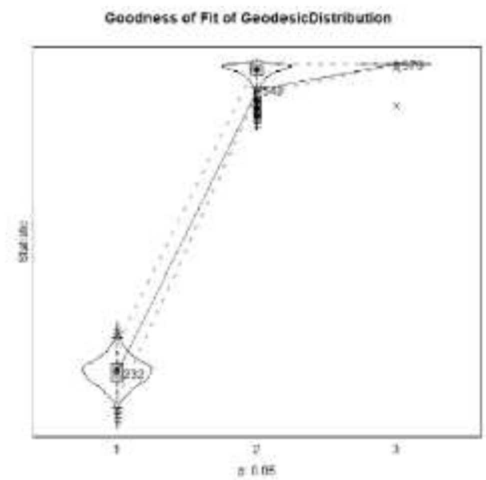
(a) Class 7



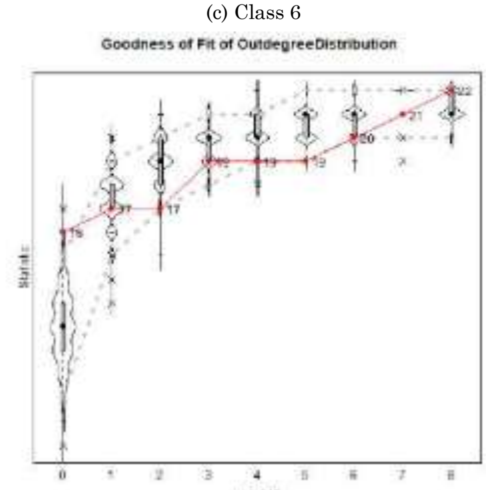
(b) Class 18



(c) Class 6



(d) Class 8



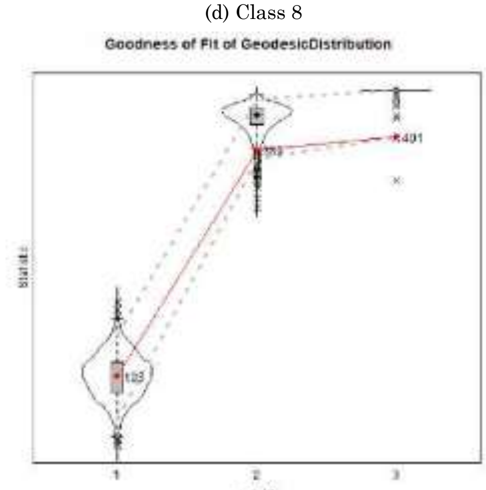


Figure 13: Classes with poor fit, Negative and TG, Model 1

(e) Class 10

(f) Class 13

Table 40: Results of separate SAOMs for negative network and TG, Model 1

	1	2	3	4	5	6	7	8	9	10
outdegree (density)	-1.894	1.612	-0.043	-2.765	-2.4	-1.139	-2.296	-1.713	-1.52	0.757
	[0.4]	[8.228]	[1.652]	[0.794]	[0.615]	[0.825]	[0.622]	[0.365]	[0.395]	[3.366]
reciprocity	0.722	2.182	0.054	0.351	1.078	0.681	1.15	0.533	0.32	0.269
	[0.229]	[1.348]	[0.491]	[0.548]	[0.315]	[0.486]	[0.367]	[0.219]	[0.272]	[1.325]
transitive triplets	-0.067	1.054	-0.163	0.354	-0.207	0.021	-0.125	-0.236	0.042	-0.045
	[0.055]	[1.142]	[0.168]	[0.241]	[0.084]	[0.165]	[0.084]	[0.069]	[0.067]	[0.658]
indegree - popularity	0.107	-0.724	-0.403	0.058	0.103	0.097	0.085	0.104	0.027	-3.027
	[0.042]	[0.638]	[0.288]	[0.117]	[0.07]	[0.047]	[0.072]	[0.042]	[0.05]	[1.984]
outdegree - activity	0.065	-1.023	0.167	0.064	0.155	0.046	0.13	0.089	0.033	0.312
	[0.02]	[2.418]	[0.061]	[0.078]	[0.031]	[0.049]	[0.031]	[0.022]	[0.025]	[0.16]
indegree - activity	0	0	0	0	0	-0.172	0	0	0	0
	[NA]	[NA]	[NA]	[NA]	[NA]	[0.146]	[NA]	[NA]	[NA]	[NA]
outdegree-trunc(1)	0	0	0	1.495	0	0	0	-2.331	0	0
	[NA]	[NA]	[NA]	[2.244]	[NA]	[NA]	[NA]	[1.385]	[NA]	[NA]
sex alter	0.053	-2.211	-2.312	0.265	-0.52	0.062	-0.034	0.274	0.234	-0.306
	[0.17]	[1.737]	[0.882]	[0.484]	[0.426]	[0.244]	[0.237]	[0.204]	[0.224]	[1.238]
sex ego	-0.107	1.297	-0.358	1.286	-0.564	-0.23	-0.162	0.251	0.349	-0.651
	[0.17]	[3.415]	[0.354]	[0.622]	[0.408]	[0.299]	[0.251]	[0.177]	[0.234]	[0.929]
same sex	-0.516	-1.35	-0.488	-0.031	-1.493	-0.125	-0.445	-0.634	-0.501	-1.49
	[0.158]	[1.044]	[0.485]	[0.405]	[0.45]	[0.193]	[0.212]	[0.168]	[0.204]	[1.001]
grade alter	0.133	-0.612	0.083	0.149	0.112	0.064	0.156	-0.237	-0.262	0.512
	[0.155]	[0.73]	[0.243]	[0.216]	[0.165]	[0.105]	[0.231]	[0.115]	[0.148]	[0.747]
grade ego	0.083	-0.104	-0.048	-0.135	0.088	0.08	0.54	-0.07	-0.184	-0.365
	[0.151]	[1.838]	[0.183]	[0.233]	[0.147]	[0.1]	[0.263]	[0.108]	[0.151]	[0.905]
grade similarity	0.004	-0.749	0.75	-0.305	-0.256	-0.324	-0.324	-0.591	0.158	4.056
	[0.275]	[2.208]	[0.497]	[0.808]	[0.863]	[0.453]	[0.345]	[0.306]	[0.29]	[2]
TG alter	0.027	-1.051	0.198	0.13	-0.027	-0.151	-0.227	-0.093	-0.041	1.735
	[0.037]	[0.837]	[0.161]	[0.08]	[0.048]	[0.076]	[0.141]	[0.06]	[0.067]	[0.982]
TG ego	-0.117	-0.732	-0.211	0.186	-0.128	-0.061	0.055	0.115	-0.019	-0.381
	[0.048]	[1.043]	[0.104]	[0.113]	[0.052]	[0.089]	[0.145]	[0.056]	[0.074]	[0.359]
TG similarity	0.062	-7.4	-0.489	0.762	-0.776	-0.002	1.108	-0.344	-0.218	3.587
	[0.405]	[6.297]	[0.742]	[0.864]	[0.462]	[0.411]	[0.533]	[0.34]	[0.384]	[2.029]
popularity alter	-0.087	-0.361	-0.991	-0.32	-0.244	-0.08	-0.335	-0.174	-0.124	-2.515
	[0.061]	[0.601]	[0.447]	[0.166]	[0.102]	[0.077]	[0.151]	[0.061]	[0.084]	[1.262]
popularity ego	0.014	-2.086	-0.099	-0.002	-0.055	-0.222	-0.001	-0.031	-0.096	-0.04
	[0.035]	[5.215]	[0.103]	[0.129]	[0.068]	[0.183]	[0.09]	[0.04]	[0.058]	[0.496]
popularity ego x popularity alter	-0.021	0.16	0.029	-0.119	0.02	-0.017	0.014	-0.019	0.007	0.073
	[0.009]	[0.355]	[0.036]	[0.053]	[0.035]	[0.02]	[0.038]	[0.015]	[0.017]	[0.123]
Overall maximum convergence ratio:	0.138	0.168	0.139	0.152	0.137	0.165	0.172	0.144	0.128	0.208

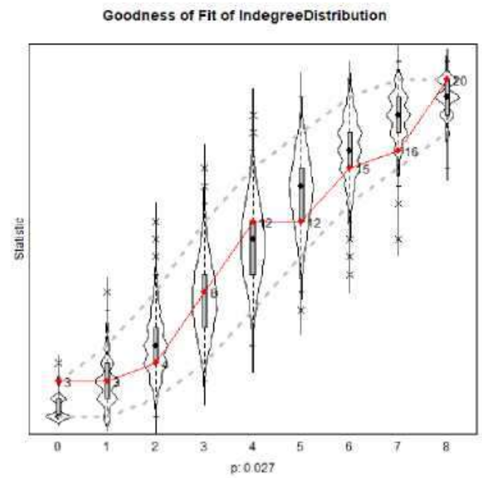
	11	12	13	14	15	16	17	18	19	20
outdegree (density)	-1.746	-2.872	-1.576	-2.583	-3.112	-2.756	-2.834	-2.421	-0.37	-1.585
	[0.618]	[0.538]	[0.942]	[0.492]	[0.512]	[1.211]	[0.928]	[0.455]	[0.994]	[0.921]
reciprocity	0.466	0.959	0.809	0.578	-0.089	0.451	0.256	0.528	-0.96	0.152
	[0.368]	[0.328]	[0.341]	[0.277]	[0.392]	[0.306]	[0.797]	[0.42]	[1.062]	[0.598]
transitive triplets	-0.251	-0.042	-0.004	-0.113	-0.187	-0.124	-0.91	-0.08	-0.548	-0.288
	[0.093]	[0.061]	[0.122]	[0.062]	[0.144]	[0.105]	[0.932]	[0.097]	[0.541]	[0.232]
indegree - popularity	0.111	0.099	0.097	0.154	0.186	0.11	0.005	0.174	-0.05	-0.018
	[0.066]	[0.063]	[0.061]	[0.043]	[0.054]	[0.062]	[0.181]	[0.044]	[0.186]	[0.207]
outdegree - activity	0.107	0.099	0.065	0.095	0.165	0.098	0.052	0.098	0.164	0.211
	[0.036]	[0.024]	[0.041]	[0.027]	[0.04]	[0.031]	[0.066]	[0.031]	[0.123]	[0.066]
indegree - activity	0	0	-0.073	0	0	0.061	0	0.023	0	0
	[NA]	[NA]	[0.113]	[NA]	[NA]	[0.128]	[NA]	[0.06]	[NA]	[NA]
outdegree-trunc(1)	-4.478	0	0	0	0	0	0	0	-4.335	0
	[1.38]	[NA]	[NA]	[NA]	[NA]	[NA]	[NA]	[NA]	[1.635]	[NA]
sex alter	-0.335	-0.257	-0.125	0.29	1.19	0.067	-1.121	0.17	-1.813	-0.785
	[0.224]	[0.252]	[0.308]	[0.197]	[0.356]	[0.231]	[0.708]	[0.234]	[1.506]	[0.518]
sex ego	0.01	0.19	0.258	0.301	0.802	0.512	2.452	0.486	-1.02	-0.032
	[0.228]	[0.222]	[0.277]	[0.199]	[0.306]	[0.352]	[0.954]	[0.25]	[0.796]	[0.377]
same sex	-0.247	-0.272	-0.428	-0.482	-1.063	-0.507	0.575	-1.03	-1.498	-1.481
	[0.222]	[0.212]	[0.285]	[0.201]	[0.327]	[0.191]	[0.585]	[0.301]	[0.706]	[0.514]
grade alter	0.059	-0.011	0.126	-0.038	0.041	0.03	0.117	-0.095	0.572	0.269
	[0.123]	[0.198]	[0.115]	[0.108]	[0.123]	[0.118]	[0.194]	[0.121]	[0.465]	[0.236]
grade ego	-0.003	-0.122	-0.135	0.013	-0.07	0.186	0.039	0.011	0.049	0.014
	[0.131]	[0.19]	[0.1]	[0.125]	[0.117]	[0.117]	[0.197]	[0.123]	[0.247]	[0.195]
grade similarity	-0.399	-0.563	-0.2	0.794	1.161	0.703	0.85	-0.136	-1.025	0.053
	[0.43]	[0.418]	[0.432]	[0.463]	[0.62]	[0.393]	[1.027]	[0.42]	[1.005]	[0.703]
TG alter	0.077	-0.04	0.044	0.095	0.069	-0.042	-0.133	-0.116	-0.088	0.163
	[0.048]	[0.05]	[0.063]	[0.063]	[0.058]	[0.067]	[0.091]	[0.059]	[0.136]	[0.133]
TG ego	0.032	-0.036	0.067	0.142	-0.106	0.113	0.075	-0.014	-0.041	-0.086
	[0.046]	[0.053]	[0.069]	[0.07]	[0.06]	[0.105]	[0.111]	[0.056]	[0.084]	[0.113]
TG similarity	0.146	-0.341	-0.575	-0.134	-0.283	-0.026	-1.952	-0.053	-0.408	0.187
	[0.502]	[0.36]	[0.406]	[0.352]	[0.707]	[0.3]	[0.795]	[0.592]	[0.717]	[0.646]
popularity alter	-0.299	-0.121	-0.13	-0.054	-0.042	-0.255	-0.267	-0.08	-0.778	-0.415
	[0.115]	[0.06]	[0.106]	[0.067]	[0.109]	[0.102]	[0.155]	[0.087]	[0.39]	[0.195]
popularity ego	-0.049	0.016	0.034	0.044	-0.026	0.193	-0.114	0.126	-0.015	-0.061
	[0.067]	[0.042]	[0.157]	[0.056]	[0.084]	[0.27]	[0.108]	[0.129]	[0.124]	[0.098]
popularity ego x popularity alter	-0.022	-0.011	-0.016	-0.03	-0.024	0.027	-0.041	-0.019	-0.024	-0.051
	[0.02]	[0.013]	[0.03]	[0.023]	[0.034]	[0.026]	[0.07]	[0.039]	[0.043]	[0.041]
Overall maximum convergence ratio:	0.112	0.095	0.164	0.118	0.140	0.147	0.149	0.189	0.190	0.144

Table 41: Results of the meta-analysis for negative network and TG, Model 2

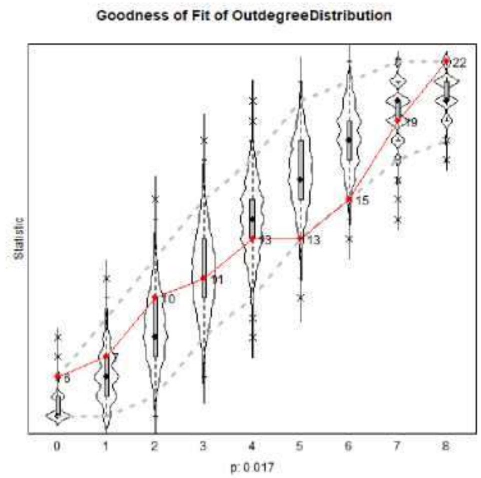
	est	se	N	p	tau2	Q	Qp
outdegree (density)	-2.131	0.1539	20	0.000	0.218	19.339	0.435
reciprocity	0.582	0.0810	20	0.000	0.000	16.496	0.624
transitive triplets	-0.107	0.0271	20	0.000	0.054	23.497	0.216
indegree - popularity	0.112	0.0145	20	0.000	0.001	15.985	0.658
outdegree - activity	0.098	0.0102	20	0.000	0.024	27.767	0.088
indegree - activity	-0.004	0.0480	4	0.926	0.000	2.206	0.531
outdegree-trunc(1)	-2.834	1.0569	4	0.007	1.343	5.641	0.130
sex alter	0.027	0.0818	20	0.745	0.182	35.391	0.013
sex ego	0.165	0.0762	20	0.030	0.147	31.195	0.038
same sex	-0.507	0.0641	20	0.000	0.097	29.730	0.055
grade alter	0.011	0.0340	20	0.745	0.000	16.505	0.623
grade ego	0.001	0.0334	20	0.965	0.000	11.489	0.906
grade similarity	0.004	0.1253	20	0.974	0.264	25.781	0.136
TG alter	0.015	0.0205	20	0.479	0.012	19.952	0.397
TG ego	-0.021	0.0252	20	0.403	0.068	32.162	0.030
TG similarity	-0.126	0.1151	20	0.272	0.001	19.912	0.400
popularity alter	-0.144	0.0221	20	0.000	0.001	23.530	0.215
popularity ego	-0.017	0.0167	20	0.316	0.000	9.772	0.958
popularity ego x popularity alter	-0.015	0.0049	20	0.002	0.000	14.188	0.773
int. TG ego x TG similarity	-0.087	0.0596	20	0.144	0.000	20.224	0.381

Table 42: Goodness of fit statistics, Negative and TG, Model 2

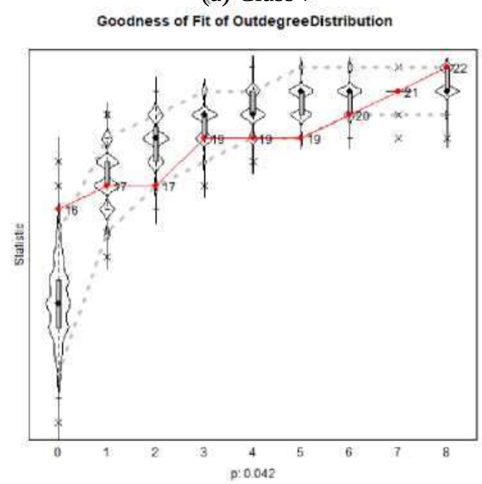
	1	2	3	4
1	0.790	0.579	0.962	0.656
2	0.926	0.925	0.991	0.783
3	0.725	0.956	1.000	0.719
4	0.300	0.332	0.086	0.117
5	0.321	0.129	0.054	0.977
6	0.973	0.017	0.010	0.998
7	0.027	0.812	0.335	0.840
8	0.944	0.596	0.044	0.976
9	0.864	0.498	0.088	0.915
10	0.500	0.042	0.648	0.852
11	0.688	0.736	0.870	0.756
12	0.167	0.497	0.907	0.992
13	0.991	0.172	0.029	0.517
14	0.376	0.128	0.349	0.938
15	0.988	0.265	0.525	0.310
16	0.796	0.834	0.085	0.760
17	0.996	0.822	0.959	1.000
18	0.053	0.307	0.537	0.414
19	0.986	0.753	0.208	0.689
20	0.448	0.602	0.937	0.841



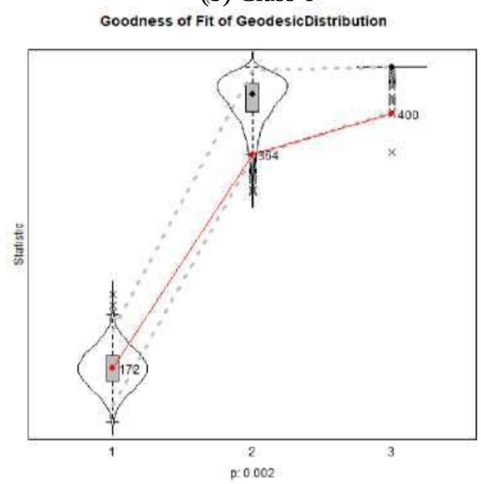
(a) Class 7



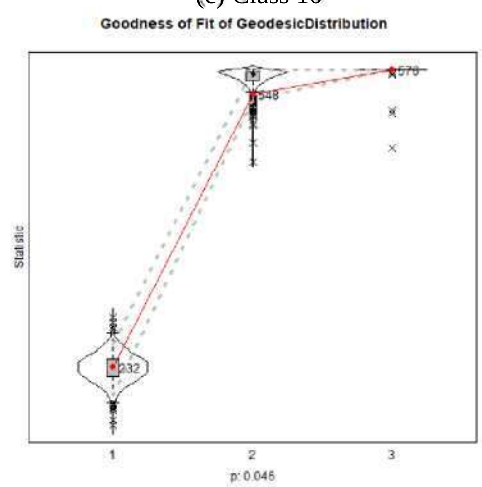
(b) Class 6



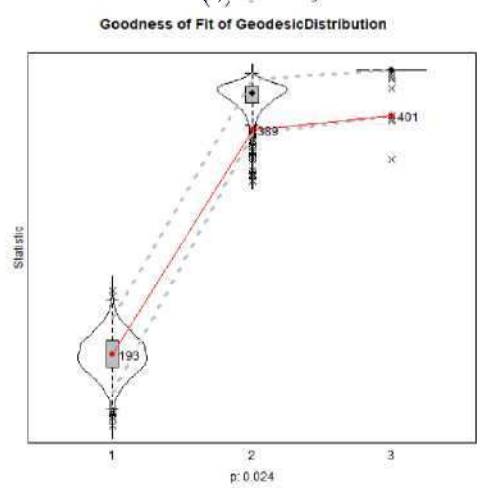
(c) Class 10



(d) Class 6



(e) Class 8



(f) Class 13

Figure 14: Classes with poor fit, Negative and TG, Model 2

Table 43: Results of separate SAOMs for negative network and TG, Model 2

	1	2	3	4	5	6	7	8	9	10
outdegree (density)	-1.901	3.088	-0.073	-2.719	-2.435	-1.304	-2.499	-1.688	-1.526	0.473
	[0.398]	[6.844]	[1.595]	[0.837]	[0.616]	[0.852]	[0.735]	[0.376]	[0.403]	[3.504]
reciprocity	0.715	2.756	0.12	0.327	1.07	0.672	1.201	0.528	0.314	0.285
	[0.229]	[1.974]	[0.47]	[0.527]	[0.302]	[0.465]	[0.416]	[0.22]	[0.276]	[1.329]
transitive triplets	-0.065	1.31	-0.185	0.358	-0.214	0.006	-0.14	-0.239	0.036	-0.026
	[0.055]	[1.282]	[0.166]	[0.239]	[0.084]	[0.142]	[0.091]	[0.071]	[0.062]	[0.76]
indegree - popularity	0.106	-0.979	-0.402	0.058	0.102	0.097	0.1	0.105	0.029	-3.062
	[0.04]	[0.867]	[0.279]	[0.114]	[0.07]	[0.049]	[0.077]	[0.043]	[0.053]	[2.069]
outdegree - activity	0.064	-1.8	0.17	0.058	0.157	0.052	0.145	0.087	0.034	0.329
	[0.021]	[2.556]	[0.055]	[0.075]	[0.03]	[0.045]	[0.039]	[0.024]	[0.024]	[0.188]
indegree - activity	0	0	0	0	0	-0.15	0	0	0	0
	[NA]	[NA]	[NA]	[NA]	[NA]	[0.134]	[NA]	[NA]	[NA]	[NA]
outdegree-trunc(1)	0	0	0	1.535	0	0	0	-2.433	0	0
	[NA]	[NA]	[NA]	[2.305]	[NA]	[NA]	[NA]	[1.305]	[NA]	[NA]
sex alter	0.058	-2.789	-2.227	0.301	-0.534	0.054	-0.067	0.281	0.241	-0.379
	[0.162]	[2.079]	[0.884]	[0.524]	[0.428]	[0.236]	[0.265]	[0.203]	[0.226]	[1.139]
sex ego	-0.113	1.855	-0.291	1.42	-0.578	-0.193	-0.187	0.262	0.331	-0.783
	[0.171]	[3.968]	[0.349]	[0.645]	[0.414]	[0.292]	[0.27]	[0.18]	[0.245]	[0.944]
same sex	-0.524	-1.256	-0.371	-0.037	-1.485	-0.134	-0.482	-0.63	-0.503	-1.362
	[0.159]	[1.086]	[0.512]	[0.433]	[0.45]	[0.191]	[0.235]	[0.169]	[0.206]	[1.067]
grade alter	0.136	-0.676	0.091	0.141	0.109	0.065	0.172	-0.236	-0.257	0.436
	[0.145]	[0.787]	[0.283]	[0.222]	[0.168]	[0.098]	[0.211]	[0.117]	[0.141]	[0.964]
grade ego	0.081	0.29	-0.113	-0.152	0.084	0.079	0.527	-0.072	-0.161	-0.416
	[0.155]	[1.399]	[0.205]	[0.236]	[0.144]	[0.1]	[0.331]	[0.106]	[0.154]	[1.092]
grade similarity	0	-0.341	0.769	-0.324	-0.213	-0.34	-0.357	-0.607	0.14	4.014
	[0.267]	[1.66]	[0.538]	[0.828]	[0.901]	[0.423]	[0.379]	[0.311]	[0.295]	[2.105]
TG alter	0.04	-0.812	0.02	0.091	0.002	-0.079	-0.02	-0.074	-0.008	1.976
	[0.049]	[0.834]	[0.185]	[0.089]	[0.07]	[0.105]	[0.161]	[0.075]	[0.085]	[1.129]
TG ego	-0.13	-1.813	-0.184	0.225	-0.141	-0.082	-0.153	0.111	-0.037	-0.517
	[0.059]	[2.049]	[0.107]	[0.132]	[0.057]	[0.091]	[0.17]	[0.058]	[0.076]	[0.437]
TG similarity	0.153	-11.072	-0.682	0.592	-0.657	0.172	1.126	-0.303	-0.196	3.651
	[0.445]	[10.907]	[0.747]	[0.874]	[0.479]	[0.46]	[0.588]	[0.357]	[0.384]	[1.971]
popularity alter	-0.089	-0.564	-0.99	-0.315	-0.246	-0.079	-0.338	-0.173	-0.123	-2.502
	[0.059]	[0.704]	[0.416]	[0.168]	[0.097]	[0.071]	[0.163]	[0.062]	[0.087]	[1.446]
popularity ego	0.013	-3.085	-0.078	-0.005	-0.052	-0.195	0.01	-0.033	-0.096	-0.024
	[0.038]	[5.149]	[0.112]	[0.136]	[0.069]	[0.181]	[0.12]	[0.042]	[0.058]	[0.579]
popularity ego x popularity alter	-0.021	0.194	0.02	-0.119	0.019	-0.017	0.024	-0.018	0.006	0.072
	[0.009]	[0.339]	[0.037]	[0.052]	[0.035]	[0.017]	[0.039]	[0.015]	[0.017]	[0.149]
int. TG ego x TG similarity	-0.059	-3.261	0.882	0.263	-0.11	-0.327	-1.456	-0.088	-0.162	-1.129
	[0.145]	[4.27]	[0.614]	[0.348]	[0.198]	[0.313]	[0.624]	[0.226]	[0.221]	[1.588]
Overall maximum convergence ratio:	0.168	0.188	0.157	0.134	0.108	0.120	0.245	0.195	0.106	0.153

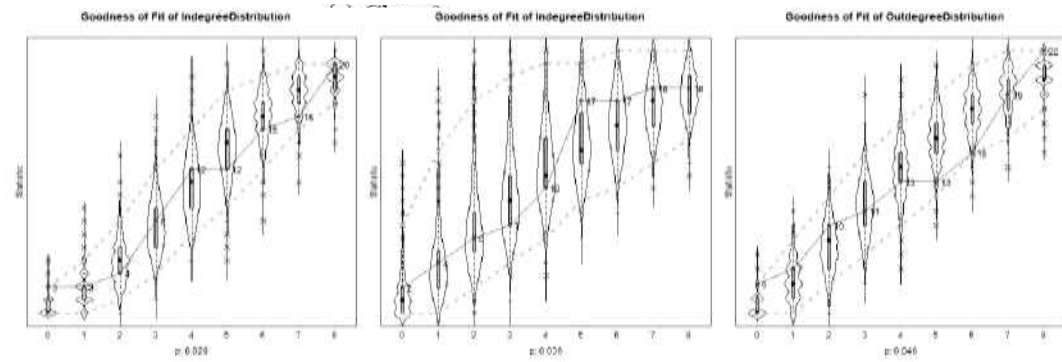
	11	12	13	14	15	16	17	18	19	20
outdegree (density)	-1.855	-2.932	-1.435	-2.584	-3.115	-2.824	-2.684	-2.536	-0.363	-1.84
	[0.628]	[0.611]	[1.14]	[0.477]	[0.556]	[1.196]	[0.818]	[0.431]	[1.111]	[0.893]
reciprocity	0.5	0.986	0.843	0.58	-0.084	0.422	0.253	0.481	-0.946	0.15
	[0.339]	[0.327]	[0.341]	[0.276]	[0.369]	[0.287]	[0.87]	[0.432]	[1.297]	[0.613]
transitive triplets	-0.25	-0.05	0.008	-0.113	-0.189	-0.124	-0.92	-0.085	-0.563	-0.31
	[0.092]	[0.07]	[0.129]	[0.059]	[0.157]	[0.097]	[0.922]	[0.099]	[0.537]	[0.22]
indegree - popularity	0.111	0.104	0.092	0.153	0.187	0.114	0.013	0.174	-0.051	0.031
	[0.067]	[0.07]	[0.065]	[0.043]	[0.061]	[0.062]	[0.159]	[0.044]	[0.19]	[0.204]
outdegree - activity	0.11	0.102	0.059	0.095	0.166	0.098	0.048	0.098	0.168	0.22
	[0.036]	[0.027]	[0.044]	[0.025]	[0.044]	[0.028]	[0.064]	[0.03]	[0.11]	[0.065]
indegree - activity			-0.092			0.07		0.028	0	
	0 [NA]	0 [NA]	[0.134]	0 [NA]	0 [NA]	[0.131]	0 [NA]	[0.062]	[NA]	0 [NA]
outdegree-trunc(1)	-4.415						0		-4.348	
	[1.454]	0 [NA]	0 [NA]	0 [NA]	0 [NA]	0 [NA]	[NA]	0 [NA]	[1.691]	0 [NA]
sex alter	-0.338	-0.274	-0.154	0.29	1.19	0.063	-1.064	0.198	-1.85	-0.662
	[0.255]	[0.245]	[0.337]	[0.195]	[0.363]	[0.225]	[0.668]	[0.248]	[1.435]	[0.535]
sex ego	0.077	0.164	0.247	0.3	0.789	0.531	2.259	0.484	-1.007	0.17
	[0.249]	[0.24]	[0.271]	[0.195]	[0.311]	[0.341]	[0.939]	[0.263]	[0.757]	[0.427]
same sex	-0.266	-0.281	-0.422	-0.474	-1.072	-0.517	0.519	-1.048	-1.502	-1.533
	[0.219]	[0.204]	[0.303]	[0.207]	[0.328]	[0.184]	[0.507]	[0.307]	[0.7]	[0.546]
grade alter	0.056	-0.017	0.131	-0.04	0.039	0.035	0.103	-0.101	0.579	0.279
	[0.122]	[0.191]	[0.127]	[0.106]	[0.13]	[0.118]	[0.196]	[0.119]	[0.506]	[0.232]
grade ego	0.033	-0.14	-0.118	0.014	-0.074	0.192	0.073	-0.019	0.059	0.063
	[0.132]	[0.191]	[0.105]	[0.113]	[0.115]	[0.118]	[0.204]	[0.133]	[0.244]	[0.203]
grade similarity	-0.393	-0.564	-0.055	0.793	1.158	0.734	0.887	-0.123	-1.015	0.044
	[0.413]	[0.419]	[0.455]	[0.457]	[0.649]	[0.385]	[1.082]	[0.41]	[1.019]	[0.737]
TG alter	0.114	-0.109	0.15	0.092	0.058	-0.078	-0.181	-0.05	-0.056	
	[0.066]	[0.107]	[0.079]	[0.095]	[0.068]	[0.101]	[0.107]	[0.076]	[0.191]	0 [0.156]
TG ego		-0.037	0.019	0.142	-0.096	0.125	0.102	-0.055	-0.049	0.024
	0 [0.058]	[0.054]	[0.089]	[0.07]	[0.071]	[0.107]	[0.127]	[0.07]	[0.091]	[0.142]
TG similarity	0.391	-0.4	-0.379	-0.121	-0.302	-0.098	-1.961	0.363	-0.391	0.109
	[0.575]	[0.387]	[0.439]	[0.355]	[0.768]	[0.331]	[0.842]	[0.693]	[0.824]	[0.736]
popularity alter	-0.296	-0.12	-0.137	-0.053	-0.041	-0.247	-0.258	-0.078	-0.785	-0.378
	[0.123]	[0.065]	[0.116]	[0.069]	[0.11]	[0.102]	[0.156]	[0.086]	[0.374]	[0.194]
popularity ego	-0.064	0.017	-0.01	0.044	-0.02	0.215	-0.131	0.137	-0.023	-0.035
	[0.069]	[0.042]	[0.193]	[0.056]	[0.081]	[0.288]	[0.114]	[0.139]	[0.123]	[0.095]
popularity ego x popularity alter	-0.024	-0.01	-0.022	-0.03	-0.024	0.029	-0.042	-0.02	-0.027	-0.054
	[0.02]	[0.014]	[0.031]	[0.023]	[0.035]	[0.028]	[0.07]	[0.04]	[0.046]	[0.041]
int. TG ego x TG similarity	-0.14	0.272	-0.641	0.004	0.107	0.123	0.301	-0.231	-0.105	0.838
	[0.156]	[0.38]	[0.308]	[0.312]	[0.311]	[0.238]	[0.429]	[0.175]	[0.422]	[0.518]
Overall maximum convergence ratio:	0.130	0.140	0.130	0.082	0.094	0.164	0.156	0.212	0.174	0.148

Table 44: Results of the meta-analysis for negative network and TGB, Model 1

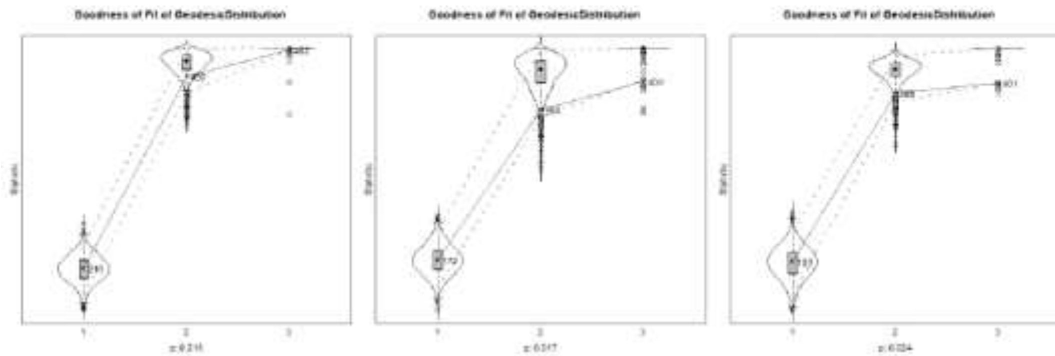
	est	se	N	p	tau2	Q	Qp
outdegree (density)	-2.110	0.1375	20	0.000	0.021	15.599	0.684
reciprocity	0.577	0.0799	20	0.000	0.002	19.787	0.407
transitive triplets	-0.096	0.0249	20	0.000	0.044	22.501	0.260
indegree - popularity	0.112	0.0137	20	0.000	0.001	14.329	0.764
outdegree - activity	0.098	0.0099	20	0.000	0.022	25.912	0.133
indegree - activity	-0.012	0.0421	4	0.785	0.000	1.189	0.756
outdegree-trunc(1)	-2.195	1.2878	4	0.088	1.984	7.211	0.065
sex alter	0.007	0.0782	20	0.929	0.169	36.713	0.009
sex ego	0.173	0.0737	19	0.019	0.137	28.562	0.054
same sex	-0.507	0.0607	20	0.000	0.068	26.224	0.124
grade alter	0.039	0.0351	20	0.264	0.037	17.590	0.550
grade ego	0.019	0.0355	20	0.597	0.035	17.817	0.535
grade similarity	-0.026	0.1119	20	0.819	0.168	24.967	0.162
TGB alter	-0.003	0.0157	20	0.844	0.001	11.878	0.891
TGB ego	-0.008	0.0264	20	0.773	0.071	33.082	0.024
TGB similarity	-0.205	0.2129	20	0.337	0.549	31.947	0.032
popularity alter	-0.145	0.0220	20	0.000	0.021	28.671	0.071
popularity ego	-0.019	0.0162	20	0.250	0.010	12.120	0.880
popularity ego x popularity alter	-0.015	0.0048	20	0.002	0.000	15.762	0.673

Table 45: Goodness of fit statistics, Negative and TGB, Model 1

	Indegree distribution	Outdegree distribution	Geodesic Distance	Triad Census
1	0.743	0.508	0.938	0.734
2	1.000	0.877	0.947	0.908
3	0.656	0.929	1.000	0.762
4	0.166	0.317	0.207	0.100
5	0.357	0.120	0.046	0.986
6	0.963	0.045	0.030	0.999
7	0.028	0.774	0.173	0.797
8	0.929	0.642	0.100	0.978
9	0.912	0.467	0.076	0.871
10	0.924	0.093	0.859	0.775
11	0.612	0.735	0.880	0.755
12	0.179	0.528	0.833	0.987
13	0.990	0.075	0.030	0.519
14	0.376	0.275	0.748	0.901
15	0.995	0.171	0.363	0.242
16	0.727	0.704	0.033	0.683
17	0.986	0.800	0.961	1.000
18	0.036	0.306	0.247	0.455
19	0.975	0.745	0.341	0.377
20	0.444	0.534	0.945	0.753



(a) Class 7 (b) Class 18 (c) Class 6



(d) Class 5 (f) Class 13 (g) Class 16

Figure 15: Classes with poor fit, Negative & TGB, Model 1

Table 46: Results of separate SAOMs for negative network and TGB, Model 1

	1	2	3	4	5	6	7	8	9	10
outdegree (density)	-1.979	1.384	-0.602	-3.14	-2.335	-1.346	-2.499	-1.905	-1.434	-2.125
	[0.346]	[6.918]	[1.439]	[1.087]	[0.605]	[1.045]	[0.684]	[0.385]	[0.405]	[3.593]
reciprocity	0.659	1.915	0.04	0.245	1.078	0.62	1.205	0.446	0.353	-0.09
	[0.223]	[1.203]	[0.497]	[0.612]	[0.303]	[0.429]	[0.406]	[0.211]	[0.261]	[1.37]
transitive triplets	-0.07	1.504	-0.2	0.312	-0.207	0.005	-0.101	-0.248	0.038	-0.049
	[0.05]	[1.545]	[0.147]	[0.246]	[0.081]	[0.175]	[0.1]	[0.071]	[0.068]	[0.629]
indegree - popularity	0.099	-0.183	-0.256	0.059	0.106	0.123	0.094	0.112	0.016	-1.319
	[0.035]	[0.46]	[0.241]	[0.131]	[0.068]	[0.043]	[0.072]	[0.043]	[0.053]	[1.279]
outdegree - activity	0.075	-1.358	0.16	0.119	0.145	0.051	0.142	0.107	0.033	0.344
	[0.017]	[2.296]	[0.048]	[0.087]	[0.029]	[0.061]	[0.045]	[0.024]	[0.025]	[0.295]
indegree - activity	0	0	0	0	0	-0.164	0	0	0	0
	[NA]	[NA]	[NA]	[NA]	[NA]	[0.17]	[NA]	[NA]	[NA]	[NA]
outdegree-trunc(1)	0	0	0	1.779	0	0	0	-2.134	0	0
	[NA]	[NA]	[NA]	[1.881]	[NA]	[NA]	[NA]	[1.297]	[NA]	[NA]
sex alter	-0.04	-3.859	-1.771	0.015	-0.442	-0.049	0.047	0.331	0.236	-0.018
	[0.153]	[3.002]	[0.661]	[0.496]	[0.402]	[0.249]	[0.232]	[0.189]	[0.217]	[2.003]
sex ego	-0.06	-0.48	-0.44	1.035	-0.376	-0.278	0.204	0.042	0.334	0
	[0.145]	[2.327]	[0.37]	[0.524]	[0.378]	[0.35]	[0.301]	[0.163]	[0.219]	[NA]
same sex	-0.515	-0.595	-0.562	-0.238	-1.388	-0.138	-0.446	-0.657	-0.528	-1.005
	[0.158]	[1.13]	[0.47]	[0.419]	[0.428]	[0.195]	[0.226]	[0.161]	[0.205]	[1.213]
grade alter	0.097	-0.224	0.031	0.622	0.15	0.023	0.339	-0.172	-0.238	0.156
	[0.14]	[0.469]	[0.245]	[0.366]	[0.158]	[0.1]	[0.171]	[0.098]	[0.15]	[0.599]
grade ego	0.106	0.654	-0.027	0.079	0.177	0.022	0.407	-0.177	-0.16	-1.627
	[0.139]	[1.083]	[0.175]	[0.256]	[0.143]	[0.114]	[0.255]	[0.104]	[0.146]	[2.292]
grade similarity	0.009	-0.546	0.721	0.251	-0.617	-0.513	-0.149	-0.639	0.103	3.563
	[0.259]	[1.781]	[0.525]	[0.863]	[0.9]	[0.46]	[0.344]	[0.316]	[0.303]	[2.093]
TGB alter	0.034	-1.767	0.02	-0.125	-0.036	0.022	-0.043	-0.023	-0.059	0.161
	[0.034]	[1.641]	[0.121]	[0.186]	[0.05]	[0.078]	[0.077]	[0.038]	[0.082]	[0.302]
TGB ego	-0.031	-2.008	-0.166	-0.154	-0.114	0.048	0.192	0.089	-0.141	0.613
	[0.035]	[2.581]	[0.094]	[0.182]	[0.057]	[0.096]	[0.135]	[0.044]	[0.087]	[0.455]
TGB similarity	0.53	-4.518	-1.082	-4.815	-0.427	1.417	0.486	0.011	-1.284	0.573
	[0.398]	[4.961]	[0.957]	[2.821]	[0.598]	[1.016]	[0.494]	[0.42]	[0.849]	[1.521]
popularity alter	-0.106	-0.016	-0.724	-0.581	-0.246	-0.057	-0.35	-0.165	-0.161	-1.605
	[0.052]	[0.42]	[0.351]	[0.233]	[0.103]	[0.07]	[0.118]	[0.061]	[0.086]	[0.979]
popularity ego	0.027	-1.544	-0.19	-0.052	-0.015	-0.202	-0.007	-0.039	-0.069	0.551
	[0.03]	[3.319]	[0.1]	[0.154]	[0.067]	[0.211]	[0.098]	[0.041]	[0.05]	[1.116]
popularity ego x popularity alter	-0.023	0.247	0.01	-0.146	0.021	-0.018	0.015	-0.02	0.008	0.077
	[0.008]	[0.319]	[0.035]	[0.066]	[0.035]	[0.018]	[0.035]	[0.016]	[0.016]	[0.14]
Overall maximum convergence ratio:	0.136	0.131	0.186	0.127	0.128	0.159	0.152	0.201	0.082	0.153

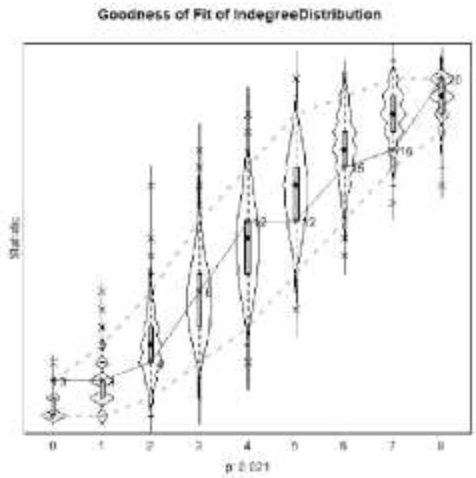
	11	12	13	14	15	16	17	18	19	20
outdegree (density)	-1.846	-2.875	-1.553	-2.482	-3.133	-2.29	-2.43	-2.315	-0.808	-1.74
	[0.65]	[0.571]	[0.907]	[0.517]	[0.536]	[1.078]	[0.871]	[0.409]	[0.913]	[0.865]
reciprocity	0.554	0.912	0.804	0.691	-0.273	0.398	0.435	0.717	-2.005	0.042
	[0.352]	[0.325]	[0.303]	[0.278]	[0.394]	[0.315]	[0.861]	[0.396]	[1.58]	[0.566]
transitive triplets	-0.222	-0.038	-0.005	-0.076	-0.217	-0.104	-0.871	-0.036	-0.802	-0.307
	[0.107]	[0.069]	[0.112]	[0.052]	[0.158]	[0.081]	[0.843]	[0.088]	[0.552]	[0.227]
indegree - popularity	0.102	0.108	0.081	0.133	0.192	0.105	-0.071	0.177	0	0.05
	[0.073]	[0.066]	[0.059]	[0.045]	[0.055]	[0.053]	[0.189]	[0.04]	[0.213]	[0.191]
outdegree - activity	0.104	0.087	0.063	0.086	0.176	0.108	0.013	0.076	0.246	0.209
	[0.041]	[0.026]	[0.036]	[0.024]	[0.045]	[0.025]	[0.093]	[0.034]	[0.113]	[0.065]
indegree - activity	0	0	-0.059	0	0	-0.025	0	0.012	0	0
	[NA]	[NA]	[0.122]	[NA]	[NA]	[0.12]	[NA]	[0.05]	[NA]	[NA]
outdegree-trunc(1)	-4.549	0	0	0	0	0	0	0	-3.479	0
	[1.553]	[NA]	[NA]	[NA]	[NA]	[NA]	[NA]	[NA]	[1.884]	[NA]
sex alter	-0.216	-0.327	-0.319	0.168	1.265	0.049	-1.205	0.27	-2.942	-0.656
	[0.238]	[0.289]	[0.287]	[0.195]	[0.404]	[0.228]	[0.847]	[0.211]	[1.606]	[0.52]
sex ego	0.103	0.514	0.078	0.343	0.931	0.463	1.928	0.364	-2.025	-0.171
	[0.247]	[0.322]	[0.265]	[0.198]	[0.348]	[0.251]	[0.996]	[0.235]	[1.294]	[0.394]
same sex	-0.235	-0.23	-0.432	-0.487	-1.181	-0.501	0.47	-0.937	-1.842	-1.473
	[0.218]	[0.222]	[0.245]	[0.201]	[0.364]	[0.196]	[0.625]	[0.287]	[0.919]	[0.525]
grade alter	0.124	-0.006	0.151	0.045	0.015	0.033	-0.042	0.018	0.441	0.249
	[0.128]	[0.212]	[0.114]	[0.101]	[0.132]	[0.122]	[0.237]	[0.13]	[0.481]	[0.28]
grade ego	0.051	-0.319	-0.103	0.111	-0.027	0.169	0.357	0.05	-0.198	0.181
	[0.136]	[0.225]	[0.127]	[0.106]	[0.113]	[0.115]	[0.369]	[0.137]	[0.28]	[0.237]
grade similarity	-0.411	-0.546	-0.195	0.539	1.326	0.435	1.413	-0.101	-1.554	0.272
	[0.424]	[0.418]	[0.447]	[0.427]	[0.636]	[0.392]	[1.045]	[0.39]	[1.113]	[0.67]
TGB alter	0.143	-0.038	0.066	-0.043	0.042	-0.061	-0.363	0.017	0.076	0.05
	[0.133]	[0.071]	[0.063]	[0.057]	[0.078]	[0.072]	[0.223]	[0.069]	[0.154]	[0.147]
TGB ego	0.102	-0.166	0.032	0.096	-0.071	-0.144	0.121	0.085	0.107	0.1
	[0.12]	[0.084]	[0.079]	[0.075]	[0.075]	[0.088]	[0.194]	[0.073]	[0.108]	[0.124]
TGB similarity	0.886	-0.862	0.892	-1.566	-0.883	-0.335	-1.351	0.867	-2.113	0.467
	[1.417]	[0.662]	[0.576]	[0.671]	[0.424]	[0.64]	[1.291]	[0.831]	[1.193]	[0.798]
popularity alter	-0.308	-0.098	-0.16	-0.077	0.027	-0.262	-0.27	-0.045	-0.834	-0.405
	[0.123]	[0.06]	[0.103]	[0.065]	[0.128]	[0.101]	[0.179]	[0.074]	[0.394]	[0.189]
popularity ego	-0.052	-0.009	0.037	0.016	-0.038	-0.011	-0.198	0.082	0.064	-0.106
	[0.065]	[0.048]	[0.189]	[0.053]	[0.095]	[0.237]	[0.142]	[0.101]	[0.151]	[0.113]
popularity ego x popularity alter	-0.026	-0.012	-0.012	-0.03	0.005	0.035	-0.04	-0.018	-0.016	-0.055
	[0.022]	[0.012]	[0.032]	[0.022]	[0.039]	[0.028]	[0.072]	[0.036]	[0.049]	[0.04]
Overall maximum convergence ratio:	0.214	0.157	0.159	0.126	0.134	0.137	0.154	0.161	0.212	0.128

Table 47: Results of the meta-analysis for negative network and TGB, Model 2

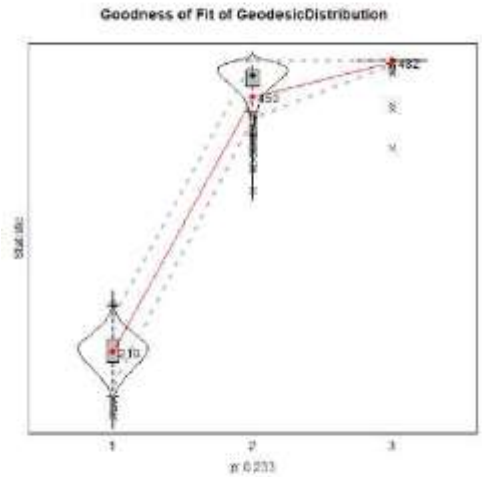
	est	se	N	p	tau2	Q	Qp
outdegree (density)	-2.128	0.142	20	0.000	0.000	13.236	0.826
reciprocity	0.591	0.080	20	0.000	0.001	19.624	0.418
transitive triplets	-0.095	0.024	20	0.000	0.040	22.175	0.276
indegree - popularity	0.114	0.014	20	0.000	0.002	14.281	0.767
outdegree - activity	0.098	0.010	20	0.000	0.021	26.164	0.126
indegree - activity	-0.019	0.045	4	0.672	0.000	2.074	0.557
outdegree-trunc(1)	-2.199	1.361	4	0.106	2.091	7.333	0.062
sex alter	0.024	0.077	20	0.756	0.162	36.170	0.010
sex ego	0.178	0.078	19	0.023	0.162	30.996	0.029
same sex	-0.510	0.067	20	0.000	0.117	28.613	0.072
grade alter	0.034	0.035	20	0.333	0.016	17.005	0.590
grade ego	0.029	0.039	20	0.465	0.062	19.492	0.426
grade similarity	-0.045	0.112	20	0.691	0.172	24.937	0.163
TGB alter	-0.014	0.018	20	0.433	0.000	12.943	0.841
TGB ego	0.000	0.036	20	0.998	0.108	40.943	0.002
TGB similarity	-0.295	0.217	20	0.174	0.547	29.831	0.054
popularity alter	-0.147	0.024	20	0.000	0.037	29.148	0.064
popularity ego	-0.015	0.016	20	0.344	0.000	10.104	0.950
popularity ego x popularity alter	-0.016	0.005	20	0.001	0.000	15.893	0.664
int. TGB ego x TGB similarity	0.037	0.066	20	0.577	0.079	19.797	0.407

Table 48: Goodness of fit statistics, Negative and TGB, Model 2

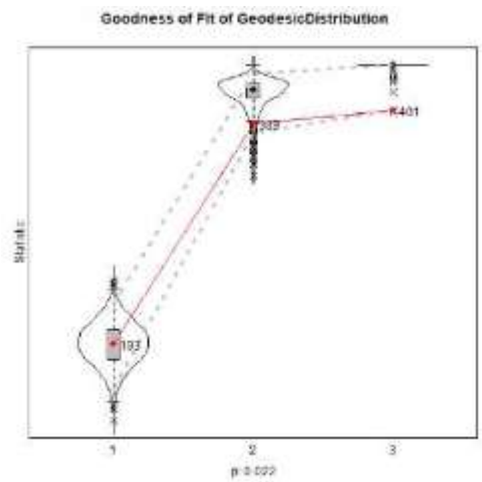
	Indegree distribution	Outdegree distribution	Geodesic Distance	Triad Census
1	0.750	0.547	0.938	0.810
2	1.000	0.901	0.931	0.935
3	0.639	0.954	1.000	0.815
4	0.224	0.280	0.148	0.130
5	0.341	0.096	0.040	0.985
6	0.976	0.157	0.114	1.000
7	0.021	0.763	0.185	0.785
8	0.952	0.641	0.106	0.980
9	0.923	0.420	0.086	0.871
10	0.946	0.075	0.824	0.769
11	0.727	0.792	0.879	0.805
12	0.227	0.478	0.843	0.972
13	0.987	0.088	0.034	0.499
14	0.411	0.246	0.791	0.861
15	0.989	0.131	0.440	0.318
16	0.649	0.620	0.026	0.672
17	0.994	0.839	0.948	1.000
18	0.058	0.250	0.140	0.444
19	0.984	0.748	0.433	0.501
20	0.470	0.562	0.950	0.725



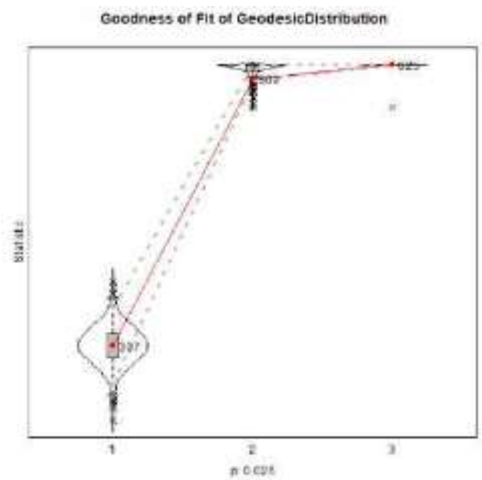
(a) Class 7



(b) Class 5



(c) Class 13



(d) Class 16

Figure 16: Classes with poor fit, Negative and TGB, Model 2

Table 271: Results of separate SAOMs for negative network and TGB, Model 2

	1	2	3	4	5	6	7	8	9	10
outdegree (density)	-1.969	1.26	-0.771	-3.019	-2.409	-0.961	-2.545	-1.936	-1.548	-1.895
	[0.362]	[6.373]	[1.629]	[1.029]	[0.622]	[1.064]	[0.635]	[0.374]	[0.445]	[3.031]
reciprocity	0.652	1.865	0.08	0.233	1.078	0.646	1.243	0.454	0.352	0.045
	[0.22]	[1.222]	[0.469]	[0.622]	[0.299]	[0.413]	[0.388]	[0.214]	[0.275]	[1.339]
transitive triplets	-0.07	1.563	-0.196	0.324	-0.203	0	-0.109	-0.252	0.029	-0.028
	[0.048]	[1.645]	[0.155]	[0.251]	[0.079]	[0.185]	[0.089]	[0.073]	[0.069]	[0.727]
indegree - popularity	0.101	-0.21	-0.251	0.048	0.103	0.134	0.096	0.115	0.023	-1.347
	[0.035]	[0.481]	[0.28]	[0.13]	[0.074]	[0.043]	[0.065]	[0.043]	[0.055]	[1.045]
outdegree - activity	0.075	-1.305	0.154	0.114	0.148	0.03	0.145	0.109	0.03	0.339
	[0.018]	[2.101]	[0.05]	[0.087]	[0.03]	[0.067]	[0.042]	[0.023]	[0.026]	[0.272]
indegree - activity	0	0	0	0	0	-0.197	0	0	0	0
	[NA]	[NA]	[NA]	[NA]	[NA]	[0.152]	[NA]	[NA]	[NA]	[NA]
outdegree-trunc(1)	0	0	0	1.722	0	0	0	-2.027	0	0
	[NA]	[NA]	[NA]	[1.87]	[NA]	[NA]	[NA]	[1.545]	[NA]	[NA]
sex alter	-0.034	-3.898	-1.829	-0.006	-0.492	-0.001	0.04	0.345	0.252	0.121
	[0.156]	[3.079]	[0.769]	[0.486]	[0.393]	[0.235]	[0.227]	[0.193]	[0.23]	[1.746]
sex ego	-0.071	-0.424	-0.555	0.999	-0.387	-0.359	0.228	0.047	0.379	0
	[0.155]	[2.252]	[0.386]	[0.538]	[0.384]	[0.377]	[0.324]	[0.167]	[0.226]	[NA]
same sex	-0.51	-0.558	-0.602	-0.235	-1.436	-0.094	-0.462	-0.659	-0.522	-1.186
	[0.148]	[1.304]	[0.457]	[0.411]	[0.436]	[0.208]	[0.233]	[0.175]	[0.216]	[1.174]
grade alter	0.101	-0.215	0.016	0.607	0.157	0.048	0.364	-0.17	-0.238	0.136
	[0.142]	[0.474]	[0.245]	[0.345]	[0.17]	[0.106]	[0.196]	[0.1]	[0.152]	[0.642]
grade ego	0.104	0.61	-0.018	0.071	0.192	0.084	0.433	-0.184	-0.143	-1.74
	[0.14]	[1.062]	[0.191]	[0.247]	[0.153]	[0.135]	[0.268]	[0.101]	[0.15]	[2.124]
grade similarity	-0.006	-0.563	0.779	0.252	-0.699	-0.583	-0.162	-0.615	0.06	3.54
	[0.248]	[1.693]	[0.525]	[0.916]	[0.938]	[0.467]	[0.36]	[0.29]	[0.303]	[2.102]
TGB alter	0.019	-1.822	0.107	-0.138	0.018	-0.093	-0.116	-0.035	-0.011	-0.04
	[0.039]	[1.834]	[0.133]	[0.178]	[0.061]	[0.086]	[0.111]	[0.045]	[0.094]	[0.38]
TGB ego	-0.013	-1.941	-0.237	-0.126	-0.186	0.273	0.24	0.104	-0.238	0.75
	[0.044]	[2.429]	[0.107]	[0.215]	[0.076]	[0.163]	[0.148]	[0.048]	[0.12]	[0.564]
TGB similarity	0.447	-4.619	-0.752	-4.833	-0.215	0.464	0.575	-0.061	-1.131	-1.249
	[0.404]	[5.106]	[0.959]	[2.537]	[0.62]	[1.06]	[0.53]	[0.457]	[0.92]	[2.901]
popularity alter	-0.103	-0.017	-0.722	-0.576	-0.254	-0.049	-0.36	-0.166	-0.157	-1.72
	[0.053]	[0.426]	[0.378]	[0.242]	[0.105]	[0.073]	[0.122]	[0.058]	[0.088]	[1.001]
popularity ego	0.027	-1.449	-0.151	-0.054	-0.019	-0.144	-0.002	-0.038	-0.053	0.597
	[0.03]	[3.337]	[0.107]	[0.149]	[0.07]	[0.155]	[0.099]	[0.041]	[0.053]	[0.943]
popularity ego x popularity alter	-0.023	0.258	0.023	-0.146	0.028	-0.018	0.016	-0.019	0.005	0.053
	[0.008]	[0.342]	[0.039]	[0.067]	[0.036]	[0.021]	[0.037]	[0.016]	[0.017]	[0.11]
int. TGB ego x TGB similarity	0.101	0.161	-0.455	0.09	-0.453	0.847	0.466	0.092	-0.286	0.991
	[0.155]	[2.741]	[0.292]	[0.367]	[0.253]	[0.52]	[0.485]	[0.15]	[0.216]	[1.293]
Overall maximum convergence ratio:	0.126	0.089	0.190	0.130	0.113	0.180	0.143	0.128	0.130	0.159

	11	12	13	14	15	16	17	18	19	20
outdegree (density)	-1.859	-2.852	-1.596	-2.467	-3.185	-2.241	-2.444	-2.191	-0.886	-1.672
	[0.715]	[0.572]	[0.852]	[0.531]	[0.585]	[0.899]	[0.899]	[0.439]	[1.086]	[0.95]
reciprocity	0.555	0.952	0.798	0.71	-0.218	0.395	0.444	0.741	-1.951	-0.001
	[0.368]	[0.353]	[0.312]	[0.267]	[0.439]	[0.293]	[0.87]	[0.421]	[1.386]	[0.571]
transitive triplets	-0.228	-0.033	-0.012	-0.07	-0.232	-0.098	-0.841	-0.036	-0.78	-0.292
	[0.102]	[0.069]	[0.106]	[0.05]	[0.173]	[0.081]	[0.895]	[0.09]	[0.513]	[0.224]
indegree - popularity	0.105	0.107	0.083	0.129	0.203	0.104	-0.073	0.18	0.003	0.04
	[0.072]	[0.066]	[0.057]	[0.045]	[0.059]	[0.048]	[0.192]	[0.041]	[0.214]	[0.213]
outdegree - activity	0.105	0.089	0.066	0.085	0.184	0.107	0.014	0.068	0.256	0.205
	[0.04]	[0.026]	[0.034]	[0.024]	[0.051]	[0.026]	[0.087]	[0.04]	[0.123]	[0.065]
indegree - activity			-0.055			-0.036		0.025	0	
	0 [NA]	0 [NA]	[0.107]	0 [NA]	0 [NA]	[0.105]	0 [NA]	[0.06]	[NA]	0 [NA]
outdegree-trunc(1)	-4.56						0		-3.668	
	[1.507]	0 [NA]	0 [NA]	0 [NA]	0 [NA]	0 [NA]	[NA]	0 [NA]	[2.087]	0 [NA]
sex alter	-0.204	-0.318	-0.292	0.152	1.253	0.062	-1.208	0.259	-2.823	-0.642
	[0.242]	[0.288]	[0.323]	[0.197]	[0.376]	[0.219]	[0.828]	[0.207]	[1.511]	[0.504]
sex ego	0.118	0.453	0.082	0.351	0.924	0.414	1.991	0.338	-1.801	-0.19
	[0.258]	[0.282]	[0.284]	[0.209]	[0.331]	[0.229]	[0.893]	[0.235]	[1.062]	[0.41]
same sex	-0.223	-0.177	-0.421	-0.493	-1.146	-0.508	0.476	-0.938	-1.828	-1.506
	[0.221]	[0.211]	[0.254]	[0.199]	[0.342]	[0.205]	[0.633]	[0.287]	[0.774]	[0.554]
grade alter	0.122	-0.037	0.146	0.045	0.016	0.032	-0.029	-0.024	0.376	0.269
	[0.132]	[0.204]	[0.128]	[0.1]	[0.134]	[0.119]	[0.231]	[0.132]	[0.501]	[0.252]
grade ego	0.05	-0.268	-0.105	0.113	-0.035	0.227	0.402	0.134	-0.177	0.19
	[0.133]	[0.213]	[0.119]	[0.106]	[0.115]	[0.138]	[0.373]	[0.151]	[0.279]	[0.23]
grade similarity	-0.417	-0.495	-0.201	0.499	1.414	0.488	1.37	-0.107	-1.635	0.219
	[0.44]	[0.408]	[0.422]	[0.431]	[0.691]	[0.408]	[1.077]	[0.387]	[1.206]	[0.69]
TGB alter	0.137	-0.099	0.066	-0.024	-0.093	-0.033	-0.4	-0.061	0.073	0.082
	[0.126]	[0.084]	[0.073]	[0.059]	[0.115]	[0.083]	[0.237]	[0.094]	[0.156]	[0.171]
TGB ego	0.111	-0.12	0.027	0.093	-0.054	-0.181	0.155	0.212	0.121	0.076
	[0.134]	[0.078]	[0.097]	[0.076]	[0.08]	[0.097]	[0.227]	[0.134]	[0.106]	[0.137]
TGB similarity	0.842	-1.098	0.885	-1.593	-0.931	-0.151	-1.583	0.168	-2.06	0.541
	[1.305]	[0.643]	[0.552]	[0.677]	[0.434]	[0.741]	[1.357]	[1.026]	[0.996]	[0.783]
popularity alter	-0.305	-0.102	-0.154	-0.081	0.044	-0.265	-0.273	-0.025	-0.84	-0.413
	[0.124]	[0.062]	[0.112]	[0.067]	[0.125]	[0.097]	[0.184]	[0.074]	[0.447]	[0.192]
popularity ego	-0.049	-0.004	0.043	0.014	-0.059	-0.035	-0.203	0.065	0.048	-0.115
	[0.067]	[0.046]	[0.17]	[0.054]	[0.099]	[0.209]	[0.141]	[0.106]	[0.137]	[0.111]
popularity ego x popularity alter	-0.027	-0.016	-0.013	-0.03	-0.014	0.038	-0.038	-0.022	-0.011	-0.056
	[0.022]	[0.013]	[0.029]	[0.023]	[0.039]	[0.028]	[0.065]	[0.036]	[0.048]	[0.042]
int. TGB ego x TGB similarity	0.026	0.285	-0.015	-0.349	0.506	-0.192	0.124	0.353	0.181	-0.2
	[0.217]	[0.191]	[0.356]	[0.437]	[0.324]	[0.259]	[0.442]	[0.306]	[0.305]	[0.575]
Overall maximum convergence ratio:	0.183	0.103	0.156	0.138	0.176	0.087	0.171	0.121	0.218	0.113

6. SUMMARY AND CONCLUSION

In my dissertation, I seek answers to the question of how reputation can resolve a social dilemma where there is a conflict between the best outcomes for the individual and for the community. We conducted two experimental studies to analyse what mechanisms ensure the reliability of reputation in a situation where individuals are interested in distorting the reputation of their rivals. I also presented the results of another experiment, where I moved away from the original dilemma and study, in a more general context, what condition make signals reliable. In the last research, I show the empirical consequences of the theories on friendship and negative relations in schools. Overall, results suggest that a reliable reputation signal is not a sufficient condition for cooperation in the selected social dilemma. My dissertation is closer to basic research, but the results can be used, for instance, to develop reliable reputation systems.

In the second chapter, we examined the formation of reputation hierarchies that can provide a solution to the problem of cooperation. Reputational information contributes to cooperation by providing guidelines about previous group-beneficial or free-rider behaviour in social dilemma interactions. How reputation information could be credible, however, remains a puzzle. We tested two potential safeguards to ensure credibility: (i) reputation is a scarce resource and (ii) it is not earned for direct benefits. We tested these solutions in a laboratory experiment in which participants played two-person Prisoner's Dilemma games without partner selection, could observe some other interactions, and could communicate reputational information about possible opponents to each other. Reputational information clearly influenced cooperation decisions. Although cooperation was not sustained at a high level in any of the conditions, the possibility of exchanging third-party information was able to temporarily increase the level of strategic cooperation when reputation was a scarce resource, and reputational scores were directly translated into monetary benefits. We found that competition for monetary rewards or unrestricted non-monetary reputational rewards helped the reputation system to be informative. Finally, we found that high reputational scores are reinforced further as they are rewarded with positive messages, and positive gossip was leading to higher reputations.

The fact that we found only a small increase in cooperation under competition indicates a lack of evidence for the theory of competitive helping. Participants intended to develop a reputation system because, they gossiped about the observed decisions, evaluated others based on the gossip they received, and made decisions based on their evaluations. As a result, individuals' final actions were in line with their partners' previous actions but only when there were neither competition nor monetary rewards and when there was competition for monetary rewards. This result suggests that by building a reputation system we should take into account research on human motivations, because it seems that reputation remained reliable when internally driven cooperation was rewarded by symbolic reputation and monetary rewards were connected to competition, where strategic cooperation was more prevalent. The fact that not just gossip influenced reputation but reputation scores were used to spread gossip suggest a 'rich-get-richer' mechanism, that could result in that individual with good reputation will be more certain about their cooperative reputation and less uncertain about cooperation. The fact that there was no significant increase in cooperation may indicate that the incentives were not strong enough to switch the prisoner's dilemma to a stag hunt game.

In the third chapter, we examined the conditions that secures the credibility of gossip and as a result the strengthen the reliability of the reputation hierarchy. Communication about previous acts and passing on reputational information could be valuable for conditional action in cooperation problems and pose a punishment threat to defectors. It is an open question, however, what kind of mechanisms can make gossip honest and credible and reputational information reliable, especially if intense competition for reputations does not exclusively dictate passing on honest information. We proposed two mechanisms that could support the honesty and credibility of gossip under such a conflict of interest. One is the possibility of voluntary checks of received evaluative information from different sources and the other is social bonding between the sender and the receiver. We tested the efficiency of cross-checking and social bonding in a laboratory experiment where subjects played the Prisoner's Dilemma with gossip interactions. Although individuals had confidence in gossip in both conditions, we found that, overall, neither the opportunities for cross-checking nor bonding were able to maintain cooperation. Meanwhile, strong competition for reputation increased cooperation when individuals' payoffs depended greatly on their position relative to their rivals. Our results suggest that intense competition for reputation facilitates gossip functioning as an informal device promoting cooperation.

Compared to the first experiment we found evidence for the positive effect of competition on cooperation in the second experiment. Cooperation, however, did not emerge to a very high rate. We also found evidence for the negative gossiping about rivals. But people still could rely on positive messages, especially if it comes as a ‘gift’ before the PD in the bonding treatment. The lower confidence in gossip from individuals with bad reputation could also secure reliability. Contradicting information lowered the reliance on the reputation system in the cross-checking treatment.

In the fourth chapter we addressed the question of what makes signals informative in general. How and why animals and humans signal reliably is a key issue in biology and social sciences that needs to be understood to explain the evolution of communication. In situations in which the receiver needs to differentiate between low- and high-quality signallers, once a ruling paradigm, the Handicap Principle has claimed that honest signals have to be costly to produce. Subsequent game theoretical models, however, highlighted that honest signals are not necessarily costly. Honesty is maintained by the potential cost of cheating: by the different ratio of marginal benefit to marginal cost for low vs. high quality signallers, i.e., by differential trade-offs. Due to the difficulties of manipulating signal costs and benefits there is lack of empirical tests of these predictions. We presented the results of a laboratory decision-making experiment with human participants to test the role of equilibrium signal cost and signalling trade-offs for the development of honest communication. We found that the trade-off manipulation had much higher influence on the reliability of communication than the manipulation of the equilibrium cost of signal. Contrary to the predictions of the Handicap Principle, negative production cost promoted honesty at a very high level in the differential trade-off condition.

Our result support that informative signals could emerge if the potential benefit out-weights the signal cost for true signallers, but not for fake signallers, even if participants were tried to make signals informative in all treatments. We also proved that signals do not have to be costly for true signallers.

In the fifth chapter we examined the observable consequences of theories related to reputation-based partner choice in a school environment. Theory suggests that prosocial behavior persists because the costs implied are compensated by reputational and other social benefits and that pro-self actions are contrasted with the social costs of avoidance and exclusion threats. Namely, prosocial individuals are more likely to receive friendship nominations and less likely to receive negative ties than others. We test if such network dynamics occur in a unique dataset from twenty primary school classes in northern Italy. Social preferences of 420 students in grades 4 and 5 were elicited with incentivised social dilemma games, and full class-networks were traced in two subsequent occasions. We model the dynamics of friendship and negative ties in classrooms with Stochastic Actor Oriented Models, and conducted a meta-analysis of the results. Our key result is that, while we do not observe evidence of homophily in friendship nominations and being prosocial does not lead to more friendship nomination, individuals are significantly more likely to send negative tie nominations to peers with different offers in the dictator game. We find that social network dynamics support cooperation through avoidance between prosocial and selfish students rather than because prosocial individuals are rewarded with friendship. We found a counterintuitive result as well. Students who contributed more to the public goods are more likely to maintain negative relations with each other.

In this study we found no evidence of a preference for prosocial friends and or of the existence of negative ties towards selfish individuals, which indicate that there might be that it is not true (intrinsically motivated) altruism that makes reputational signals reliable for partner choice. In this study we found interesting result in connection with popularity that was measured in an indirect way. Children were asked about their perceived opinion about who is liked by many others in the class. I do not rule out the possibility that popularity will function as a signal when people are selecting their cooperation partners. Further research is needed to analyse how strong is the connection between cooperative reputation and popularity.

6.1. Limitations

I use laboratory experiment in order to test hypotheses about the relationship between gossip, reputation and cooperation. Laboratory experiments are commonly used in psychology, economics, but not in sociology (Webster and Sell 2014). Critical remarks about experimental methodology relates to the external validity of results founded in a quite artificial, controlled setting, and their generalizability beyond the laboratory context. These claims are valid, participants could behave differently in a real-life situation. The fact that participants were students from one university also limits the generalizability of the results. The artificial environment (e.g. the randomly chosen gossip targets in Chapter 3) could lead to unrealistic decisions. Communication designed in an abstract way may also have led to a failure to achieve cooperation which could have happened in real life. In an experiment it is also harder to remember the outcomes and the identity of the partners than in real life. In the second experiment we used fictious names instead of numbers to identify players and displayed the results of previous games on participants' screen. Even if we incentivized real decisions using monetary rewards, as a result of the small amount of the money, participants may not take the task seriously enough. Using post-experimental questionnaires, we analysed how seriously the tasks were taken. Moreover, the allocation of the participants is expected to be random, but randomization does not promise similar composition with absolute certainty. Additional analyses were made to assess the gender and age composition of the students. Experiments, therefore, should be treated as an intermediate or supplementary method of other respected empirical methodologies experiments. Having said that, I extended my draft dissertation with Chapter 5 which investigates the same research question in schools where we expected that children play a class-level meta game, next to what they were asked to play during the data collection.

6.2. Directions for future research

The research left many questions open that could be the subjects of further research. One direction that the dissertation does not address is the question of true or fake altruism (internally motivated or situation-dependent). In the dissertation I consider cooperation as a situational strategy, where individuals are expected to change their behaviour depending on the circumstances. Although we see heterogeneity among the participants in their initial willingness to cooperate, this dissertation is not able to examine the role of internal preferences in the development and maintenance of the meaning of reputational signals. On the one hand, because the data were collected in an environment where the meaning of reputation signals has been learnt already. We can understand what a smiley or a sad emoticon could mean. My question was under what conditions we rely on them. Although we examine the development of meaning in the signalling experiment, we do not link the results directly to the topic of cooperation. We might be able to distinguish between true or fake altruism in the last study if we consider anonymous games as it is not part of the meta-game of the class. In this study generous individuals were not more likely to be selected as friend. We might think that structural effects taken over the role of maintaining honest signalling. The question of how structural effects produce reliable reputational signals could be one direction of future research. In a pilot study, I tried to detect network formations that encourage individuals to cooperate, while their absence has the opposite effect.

12 classes from 1 school participates in this pilot research. On the one hand we mapped friendship relations in these classes, and we created various measures that provide information about the students' network positions. On the other hand, we measured prosociality with different economic games with slight modification of the games that we used in Chapter 5.

We tested whether prosocial behaviour correlated with friendship nominations. We found that incoming nominations do not show a correlation with prosocial behaviour. At a level higher than the even so-called dyadic level, we find interesting results. Students whose friends are also friends (or in other words are located in closed triads) offer more in the dictator game independently of how many friends they have in the class.

The proportion of closed triads in a small network cannot be distinguished from higher-level network positions, but we might say that the results are in line with research trends examining prosocial preferences and central positions. Most of these studies consider the relationship as a consequence of the preference for prosocial partners, but reverse causality is also possible. The greater visibility of individuals in central positions would result in greater contributions.

Another result of the pilot study strengthens the argument that network embeddedness is the cause, and not (just) the consequence of generosity. Examining friendship networks at the class level, we find that the average prosociality is higher in denser networks. This relationship is the subject of research on generalized reciprocity. According to this theory, a densely connected group can more effectively enforce social norms of helping and the reciprocation of help. The mixed results of research on centrality certainly motivate further discussion of these mechanism.

The research is also limited in its measurement of reputation. Studying different types of reputation in connection with the willingness to cooperate is crucial to better understand how the mechanism works. Few studies have been published already on this topic, but in my opinion, there is still room for further research. Related to this, the relationship between reputation and popularity, or between status and popularity should be the subject of future research.

6.3. Reliable reputation systems in practice

My dissertation is exploratory research that investigates the mechanisms that contribute to the achievement of beneficial group-level outcomes where it is not expected due to conflict. Competition for reputation can result in the realization of beneficial outcomes for the group in our everyday life. Public rankings, for instance, can help us in uncertain life situations: credit rating agencies, users' evaluations of a product or service, university rankings, or even recommendations from friends can be listed as examples. Although competition increases the chance of the manipulation of reputation, which destroys the reliability of the ranking, reputation rankings can remain informative. Based on my results, the reputation system accurately reflects future behaviour if positions in the reputation ranking is combined with positive and negative incentives in a way that it reduces the cost of cooperation at the top and increases the cost of manipulation for players at the bottom of the hierarchy, even if reputational information is partially false. The following real-life example where these findings can be utilized may seem a bit distant from the content of the dissertation, but, on the one hand, it is related to my previous work experience, and on the other hand, it reflects the wide generalizability of the results.

The institutions providing labour market services and training can be one example of reputation systems from everyday life. Ranking training or service providers is needed, for instance, when the unemployed can freely choose an institution by using a voucher. These rankings help individuals to decide where to spend their voucher. Such funding creates a competitive situation between the training institutions (Ziderman 2018) and provide beneficial outcomes like high-quality trainings and services.

Perfect information about the effectiveness of a training institution is, however, rarely available. Secondary signals indicating the effectiveness of an institution are essential to achieve the most effective results (voucher use in high-quality educational institutions). Since the state is interested in the efficient use of the resources, reputation system could be built on indicators that is difficult to manipulate by actors who provide low-quality services, and the return of high-quality training is maintained by partner choice. To produce the expected outcome, indicators should be selected carefully, because due to competition manipulation is also incentivized. Trainers with many applicants might select only those who contribute to good indicators (e.g., who have higher chance of re-employment). Indicators, and thus the reputation system, should encourage the re-placement of disadvantaged job seekers. The manipulability of the system depends to a large extent on the selection of the indicators or the ‘signal costs’. The only existing system that I am aware of is in a trial phase in Latvia (OECD 2019).

7. REFERENCES

Abeler, J., Nosenzo, D. and Raymond, C. 2019. Preferences for truth-telling. Econometrica 87, 1115–1153 .

Abeler, J., Nosenzo, D., and Raymond, C. 2019. Preferences for truth-telling. Econometrica, 87(4), 1115-1153.

Alexander RD. 1987 The biology of moral systems. New York, NY: Routledge.

Anderson, C., Hildreth, J.A.D. and Howland, L., 2015. Is the desire for status a fundamental human motive? A review of the empirical literature. Psychological bulletin, 141(3), p.574.

Anderson, Lisa R, Yana V Rodgers, and Roger R Rodriguez 2000. Cultural differences in attitudes toward bargaining". In: Economics Letters 69.1, pp. 45–54.

Ariely, D., Bracha, A. and Meier, S. 2009. Doing good or doing well? Image motivation and monetary incentives in behaving prosocially.Am. Econ. Rev. 99, 544–555 .

Askew, G. N. 2014. The elaborate plumage in peacocks is not such a drag. Journal of Experimental Biology, 217(18), 3237-3241.

Axelrod, R. 1984. The Evolution of Cooperation Basic Books, New York.

Backwell, P.R.Y., Christy, J.H., Telford, S. R., Jennions, M. D. and Passmore, N. I. 2000. Dishonest signalling in a fiddler crab. Proceedings of the Royal Society of London B 267, 719–724.

Baldassarri, D. and Grossman, G., 2013. The effect of group attachment and social position on prosocial behavior. Evidence from lab-in-the-field experiments. PloS one, 8(3), p.e58750.

Barclay P. 2010 Altruism as a courtship display: some effects of third-party generosity on audience perceptions. Br. J. Psychol. 101, 123-135.

Barclay, P. 2013. Strategies for cooperation in biological markets, especially for humans. Evol. Hum. Behav. 34, 164–175 .

Barclay, P. and Willer, R. 2007. Partner choice creates competitive altruism in humans. Proc. R. Soc. B Biol. Sci. 274, 749–753 .

Barclay, P. and Willer, R., 2007. Partner choice creates competitive altruism in humans. Proceedings of the Royal Society B: Biological Sciences, 274(1610), pp.749-753.

Barclay, P., 2004. Trustworthiness and competitive altruism can also solve the “tragedy of the commons”. Evolution and Human Behavior, 25(4), pp.209-220.

Barker, J. L., Power, E. A., Heap, S., Puurtinen, M., and Sosis, R. 2019. Content, cost, and context: A framework for understanding human signalling systems. Evolutionary Anthropology: Issues, News, and Reviews, 28(2), 86-99.

Barkow, J. H. 1992. Beneath new culture is old psychology: gossip and social stratification. In The adapted mind: Evolutionary psychology and the generation of culture, 627–637 Oxford University Press, New York.

Bauer, M., Chytilová, J. and Pertold-Gebicka, B., 2014. Parental background and other-regarding preferences in children. Experimental Economics, 17(1), pp.24-46.

Beecher, M. D. 2021. Why Are No Animal Communication Systems Simple Languages? Frontiers in Psychology, 12, 701.

Bellucci G, Molter F, Park SQ. 2019 Neural representations of honesty predict future trust behavior. Nat. Commun. 10, 1-2.

Bellucci G, Park SQ. 2020 Honesty biases trustworthiness impressions. J. Exp. Psychol. Gen. 149, 1567.

Bénabou, R. and Tirole, J. 2006. Incentives and prosocial behavior. Am. Econ. Rev. 96, 1652–1678 .

Berg, Joyce, John Dickhaut, and Kevin McCabe 1995. Trust, reciprocity, and social history". In: Games and economic behavior 10.1, pp. 122–142.

Berger, C. and Rodkin, P.C., 2012. Group influences on individual aggression and prosociality: Early adolescents who change peer affiliations. Social Development, 21(2), pp.396-413.

Bergstrom, C. T., and Lachmann, M. 1998. Signalling among relatives. III. Talk is cheap. Proceedings of the National Academy of Sciences, 95(9), 5100-5105.

Bergstrom, C. T., Számadó, S., and Lachmann, M. 2002. Separating equilibria in continuous signalling games. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 357(1427), 1595-1606.

Bicchieri C, Lev-On A. 2007 Computer-mediated communication and cooperation in social dilemmas: an experimental analysis. Politics Philos. Econ. 6, 139-168.

Bicchieri, C., 2005. The grammar of society: The nature and dynamics of social norms. Cambridge University Press.

Bicchieri, C., 2016. Norms in the wild: How to diagnose, measure, and change social norms. Oxford University Press.

Biernaskie, J. M., Perry, J. C. and Grafen, A. 2018. A general model of biological signals, from cues to handicaps. Evolution Letters. 2, 201–209.

Binzel, Christine and Dietmar Fehr 2013a. Giving and sorting among friends: Evidence from a lab-in-the-field experiment". In: Economics Letters 121.2, pp. 214–217.

Binzel, Christine and Dietmar Fehr 2013b. Social distance and trust: Experimental evidence from a slum in Cairo”. In: Journal of Development Economics 103, pp. 99–106.

Bird RB, Ready E, Power EA. 2018 The social significance of subtle signals. Nat. Hum. Behav. 2, 452-457.

Bird, R. B. and Power, E. A. 2015. Prosocial signaling and cooperation among Martu hunters. Evol. Hum. Behav. 36, 389–397 .

Bird, R. B., Ready, E. and Power, E. A. 2018. The social significance of subtle signals. Nat. Hum. Behav. 2, 452–457 .

Bird, R. B., Smith, E., and Bird, D. W. 2001. The hunting handicap: costly signalling in human foraging strategies. Behavioral Ecology and Sociobiology, 50(1), 9-19.

Blume, A., Lai, E. K., and Lim, W. 2020. Strategic information transmission: A survey of experiments and theoretical foundations. In Handbook of Experimental Game Theory. Edward Elgar Publishing.

Blume, Andreas, Douglas V Dejong, Yong-Gwan Kim, and Geoffrey B Sprinkle. 1998. “Experimental Evidence on the Evolution of Meaning of Messages in Sender-Receiver Games.” The American Economic Review. 88 (5): 1323–40.

Bocskor, Á. and Havelda, A., 2019. Status dynamics: Popularity and acceptance in an ethnically diverse Hungarian primary school sample. Intersections. East European Journal of Society and Politics, 5(4).

Bond, Lyndal et al. 2007. Social and school connectedness in early secondary school as predictors of late teenage substance use, mental health, and academic outcomes”. In: Journal of adolescent health 40.4, 357–e9.

Boone, J. L. 1998. The evolution of magnanimity. Human Nature, 9(1), 1-21.

Borgia, G. 1993. The cost of display in the non-resource-based mating system of the satin bowerbird. The American Naturalist, 141(5), 729-743.

Borgia, G. 1996. Satin bowerbird displays are not extremely costly. Animal Behaviour, 52(3), 648-650.

Bowles, S. 2008. Policies designed for self-interested citizens may undermine “the moral sentiments”: evidence from economic experiments. Science (New York, N.Y.) 320, 1605–1609 .

Bowles, S. and Gintis, H. 2004. The evolution of strong reciprocity: cooperation in heterogeneous populations. Theor. Popul. Biol. 65, 17–28 .

Bowles, S. and Polania-Reyes, S. 2012. Economic incentives and social preferences: substitutes or complements?. J. Econ. Lit. 50, 368–425 .

Boyd R, Mathew S. 2015 Third-party monitoring and sanctions aid the evolution of language. Evol. Hum. Behav. 36, 475-479.

Boyd, Robert and Peter J Richerson 1992. Punishment allows the evolution of cooperation or anything else) in sizable groups”. In: Ethology and sociobiology 13.3, pp. 171–195.

Bozoyan C, Vogt S. 2016 The impact of third-party information on trust: valence, source, and reliability. PLoS ONE 11, e0149542.

Bradbury, J.C., and Vehrencamp, S.L. 2011. Principles of Animal Communication. 2nd ed. Sinauer Assoc., Sunderland, Mass.

Brañas Garza, Pablo et al. 2010. Altruism and social integration”. In: Games and Economic Behavior 69.2, pp. 249–257.

Brañas Garza, Pablo et al. 2014. Fair and unfair punishers coexist in the Ultimatum Game”. In: Scientific reports 4.1, pp. 1–4.

Bravo, G., Squazzoni, F. and Takács, K. 2015. Intermediaries in trust: indirect reciprocity, incentives, and norms. J. Appl. Math. 2015, 1–12.

Bravo, Giangiacomo, Flaminio Squazzoni, and Riccardo Boero 2012. Trust and partner selection in social networks: An experimentally grounded model". In: Social Networks 34.4, pp. 481–492.

Brown, C., Garwood, M. P., and Williamson, J. E. 2012. It pays to cheat: tactical deception in a cephalopod social signalling system. Biology Letters, 8(5), 729-732.

Burt RS. 2005 Brokerage and closure: an introduction to social capital. New York, NY: Oxford University Press.

Burt, R.S., 1992. Structural holes Harvard university press. Cambridge, MA.

Buss, D. M. and Dedden, L. 1990. A. Derogation of competitors. J. Soc. Pers. Relatsh. 7, 395–422 .

Cai, Hongbin, and Joseph Tao-Yi Wang. 2006) "Overcommunication in Strategic Information Transmission Games." Games and Economic Behavior. 56 1): 7–36. https://doi.org/10.1016/j.geb.2005.04.001.

Camerer, C.F., 2011. Behavioral game theory: Experiments in strategic interaction. Princeton University Press.

Caprara, Gian Vittorio et al. 2000. Prosocial foundations of children's academic achievement". In: Psychological science 11.4, pp. 302–306.

Casewell, N. R., Visser, J. C., Baumann, K., Dobson, J., Han, H., Kuruppu, S., ... and Fry, B. G. 2017. The evolution of fangs, venom, and mimicry systems in blenny fishes. Current Biology, 27(8), 1184-1191.

Chen, Daniel L, Martin Schonger, and Chris Wickens 2016. oTree—An open-source platform for laboratory, online, and field experiments”. In: Journal of Behavioral and Experimental Finance 9, pp. 88–97.

Christy, J. H., and Rittschof, D. 2010. Deception in visual and chemical communication in crustaceans. In Chemical communication in crustaceans pp. 313-333. Springer, New York, NY.

Cillessen AHN and Marks PEL (2011) Conceptualizing and measuring popularity. In: Cillessen AHN, Schwartz D, and Mayeux L (eds) Popularity in the Peer System. New York, NY, US: The Guilford Press, pp. 25–56.

Clifton, S. M., Braun, R. I., and Abrams, D. M. 2016. Handicap principle implies emergence of dimorphic ornaments. Proceedings of the Royal Society B: Biological Sciences, 283(1843), 20161970.

Cohen, S., 2004. Social relationships and health. American psychologist, 59(8), p.676.

Coleman, J.S., 1994. Foundations of social theory. Harvard university press.

Columbus, S., Molho, C., Righetti, F. and Balliet, D., 2021. Interdependence and cooperation in daily life. Journal of Personality and Social Psychology, 120(3), p.626.

Columbus, Simon et al. 2021. Interdependence and cooperation in daily life.” In: Journal of Personality and Social Psychology 120.3, p. 626.

Coricelli, Giorgio, Dietmar Fehr, and Gerlinde Fellner 2004. Partner selection in public goods experiments”. In: Journal of Conflict Resolution 48.3, pp. 356–378.

Crawford, V., Sobel, J., 1982. Strategic information transmission. Econometrica. 50, 1431–1451.

Crawford, Vincent P. 1998. A Survey of Experiments on Communication via Cheap Talk," Journal of Economic Theory 78, 286-298.

Croson, Rachel and Uri Gneezy 2009. Gender differences in preferences”. In: Journal of Economic literature 47.2, pp. 448–74.

Curry, O.S., Mullins, D.A. and Whitehouse, H., 2019. Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. Current Anthropology, 60(1), pp.47-69.

Dalziell, A. H., and Welbergen, J. A. 2016. Mimicry for all modalities. Ecology Letters, 19(6), 609-619.

Darimont, C. T., Codding, B. F., and Hawkes, K. 2017. Why men trophy hunt. Biology Letters, 13(3), 20160909.

Dawes, R.M., 1980. Social dilemmas. Annual review of psychology. 31, 169–193.

De Backer CJ, Gurven M. 2006 Whispering down the lane: the economics of vicarious information transfer. Adapt. Behav. 14, 249-264.

De Quervain, Dominique J-F et al. 2004. The neural basis of altruistic punishment”. In: Science 305.5688, pp. 1254–1258.

De Waal F, Waal FB. 2007 Chimpanzee politics: power and sex among apes. New York, NY: Johns Hopkins University Press.

Deci, E. L., Koestner, R. and Ryan, R. M. 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. Psychol. Bull. 125, 627–668 (discussion 692–700).

Denant-Boemont, Laurent, David Masclet, and Charles N Noussair 2007. Punishment, counterpunishment and sanction enforcement in a social dilemma experiment”. In: Economic theory 33.1, pp. 145–167.

Densley, J. A. 2012. Street gang recruitment: Signalling, screening, and selection. Social problems, 59(3), 301-321.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., and Epstein, J. A. 1996. Lying in everyday life. Journal of personality and social psychology, 70(5), 979.

Dickhaut, J., McCabe, K., Mukherji, A., 1995) An experimental study of strategic information transmission. Econ. Theory. 6, 389–403.

Dijkstra, Jan Kornelis and Christian Berger 2018. Friendship selection and influence processes for physical aggression and prosociality: differences between single-sex and mixed-sex contexts". In: Sex roles 78.9, pp. 625–636.

Dores Cruz TDet al.2021 Gossip and reputation in everyday life. Phil. Trans. R. Soc. B 376, 20200301.

Duffy MK, Ganster DC, Pagon M. 2002 Social undermining in the workplace. Acad. Manag. J. 45, 331-351.

Dufwenberg, Martin and Georg Kirchsteiger (2004). A theory of sequential reciprocity. In: Games and economic behavior 47.2, pp. 268–298.

Dumas M, Barker JL, Power EA. 2021 When does reputation lie? Dynamic feedbacks between costly signals, social capital and social prominence. Phil. Trans. R. Soc. B 376, 20200298.

Dunbar R, Dunbar RI. 1998 Grooming, gossip, and the evolution of language. Cambridge, MA: Harvard University Press.

Dunbar RI. 1993 Coevolution of neocortical size, group size and language in humans. Behav. Brain Sci. 16, 681-694.

Dunbar RI. 2004 Gossip in evolutionary perspective. Rev. Gen. Psychol. 8, 100-110.

Dunbar RI. 2010 The social role of touch in humans and primates: behavioural function and neurobiological mechanisms. Neurosci. Biobehav. Rev. 34, 260-268.

Durkheim E. 1982[1895] Rules for the Explanation of Social Facts. In: Lukes S. (eds) The Rules of Sociological Method. Contemporary Social Theory. Palgrave, London

Durkheim, É. 1984[1893]. The Division of Labor in Society. New York: Free Press.

Eisenberg, N., Guthrie, I.K., Murphy, B.C., Shepard, S.A., Cumberland, A. and Carlo, G., 1999. Consistency and development of prosocial dispositions: A longitudinal study. Child development, 70(6), pp.1360-1372.

Elias, N. 2000[1939]. The civilizing process sociogenetic and psychogenetic investigations. 2nd Edition, Wiley-Blackwell

Ellwardt L, Steglich C, Wittek R. 2012 The co-evolution of gossip and friendship in workplace social networks. Soc. Netw. 34, 623-633.

Enquist, M. 1985. Communication during aggressive interactions with particular reference to variation in choice of behaviour. Animal Behaviour, 33(4), 1152-1161.

Evans, A.M. and Krueger, J.I., 2011. Elements of trust: Risk and perspective-taking. Journal of Experimental Social Psychology, 47(1), pp.171-177.

Everett, J. A. C., Faber, N. S. and Crockett, M. J. 2015. The influence of social preferences and reputational concerns on intergroup prosocial behaviour in gains and losses contexts. R. Soc. Open Sci. 2, 150546 .

Faris R, Ennett S. 2012 Adolescent aggression: the role of peer group status motives, peer aggression, and group characteristics. Soc. Netw. 34, 371-378.

Faris R, Felmlee D. 2011 Status struggles: network centrality and gender segregation in same- and cross-gender aggression. Am. Sociol. Rev. 76, 48-73.

Farthing, G. W. 2005. Attitudes toward heroic and nonheroic physical risk takers as mates and as friends. Evolution and Human Behavior, 26(2), 171-185.

Fehl, Katrin, Daniel J van der Post, and Dirk Semmann 2011. Co-evolution of behaviour and social network structure promotes human cooperation". In: Ecology letters 14.6, pp. 546–551.

Fehr, D. and Sutter, M. 2019. Gossip and the efficiency of interactions. Games Econ. Behav. 113, 448–460 .

Fehr, E and Gächter, S. 2002. Altruistic punishment in humans". In: Nature 415.6868, pp. 137–140.

Fehr, E. and Fischbacher, U. 2003. The nature of human altruism. Nature 425, 785–791 .

Fehr, E. and Fischbacher, U. 2004. Third-party punishment and social norms. Evol. Hum. Behav. 25, 63–87 .

Fehr, E. and Gächter S. 2000. Cooperation and punishment in public goods experiments. In: American Economic Review 90.4, pp. 980–994.

Fehrler, S. and Przepiorka, W., 2013. Charitable giving as a signal of trustworthiness: Disentangling the signaling benefits of altruistic acts. Evolution and Human Behavior, 34(2), pp.139-145.

Feinberg M, Willer R, Schultz M. 2014 Gossip and ostracism promote cooperation in groups. Psychol. Sci. 25, 656-664.

Feinberg, M. 2012. Exploring the Virtues of Gossip: The Prosocial Motivations and Functions of Reputational Information Sharing. Ph.D.thesis, UC Berkeley .

Feinberg, M., Cheng, J. T. and Willer, R. 2012. Gossip as an effective and low-cost form of punishment. Behav. Brain Sci. 35, 25 .

Feinberg, M., Wilier, R., Stellar, J. and Keltner, D. 2012. The virtues of gossip: reputational information sharing as prosocial behavior. J.Pers. Soc. Psychol. 102, 1015.

Feinberg, M., Willer, R. and Schultz, M. 2014. Gossip and ostracism promote cooperation in groups. Psychol. Sci. 25, 656–664 .

Feltovich, N., Harbaugh, R. and To, T., 2002. Too cool for school? Signalling and countersignalling. RAND Journal of Economics, pp.630-649.

Figuera, S. and Tortorella Esposito, G., 2019. Ethics and the Economy: Food for Thought from the Medieval Debate on Money. Journal of Economic Issues, 53(3), pp.703-725.

Fischbacher U, Gächter S, Fehr E. 2001 Are people conditionally cooperative? Evidence from a public goods experiment. Econ. Lett. 71, 397-404.

Fischbacher, U. 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10, 171–178

Fischbacher, U., and Föllmi-Heusi, F. 2013. Lies in disguise—an experimental study on cheating. Journal of the European Economic Association, 11(3), 525-547.

Fischbacher, U., and Gächter, S. 2010. Social preferences, beliefs, and the dynamics of free riding in public goods experiments. American Economic Review, 100, 541–556.

Fischbacher, U., Gächter, S. and Fehr, E. 2001. Are people conditionally cooperative? Evidence from a public goods experiment. Econ. Lett.71, 397–404 .

Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C.E. and Falk, A., 2007. Social comparison affects reward-related brain activity in the human ventral striatum. science, 318(5854), pp.1305-1308.

Fonseca M. A., Peters K. 2021 Is it costly to deceive? People are adept at detecting gossipers' lies but may not reward honesty. Phil. Trans. R. Soc. B 376, 20200304.

Fonseca, M. A. and Peters, K. 2018. Will any gossip do? Gossip does not need to be perfectly accurate to promote trust. Games Econ. Behav.107, 253–281 .

Foster, E. K. 2004. Research on gossip: taxonomy, methods, and future directions. Rev. Gen. Psychol. 8, 78–99 .

Frederick, D. A., and Haselton, M. G. 2007. Why is muscularity sexy? Tests of the fitness indicator hypothesis. Personality and Social Psychology Bulletin, 33(8), 1167-1183.

Frey, B. S. and Jegen, R. 2001. Motivation crowding theory. J. Econ. Surv. 15, 589–611 .

Fujisawa, M., Sakai, Y., and Kuwamura, T. 2020. The false cleanerfish relies on aggressive mimicry to bite fish fins when benthic foods are scarce in their local habitat. Scientific Reports, 10(1), 1-7.

Gächter, Simon and Benedikt Herrmann 2009. Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment". In: Philosophical Transactions of the Royal Society B: Biological Sciences 364.1518, pp. 791–806.

Gangestad, S. W., and Scheyd, G. J. 2005. The evolution of human physical attractiveness. Annu. Rev. Anthropol., 34, 523-548.

Garfield, Z.H., Schacht, R., Post, E.R., Ingram, D., Uehling, A. and Macfarlan, S.J., 2021. The content and structure of reputation domains across human societies: a view from the evolutionary social sciences. Philosophical Transactions of the Royal Society B, 376(1838), p.20200296.

Ge, E., Chen, Y., Wu, J. and Mace, R. 2019. Large-scale cooperation driven by reputation, not fear of divine punishment. R. Soc. Open Sci.6, 190991 .

Geary, D. C., Vigil, J., and Byrd-Craven, J. 2004. Evolution of human mate choice. Journal of sex research, 41(1), 27-42.

Getty, T. 1998. Handicap signalling: when fecundity and viability do not add up. Animal Behaviour, 56(1), 127-130.

Getty, T. 2006. Sexually selected signals are not similar to sports handicaps. Trends in Ecology and Evolution, 21(2), 83-88.

Giardini F, Conte R. 2012 Gossip for social control in natural and artificial societies. Simulation 88, 18-32.

Giardini F, Vilone D, Sánchez A, Antonioni A. 2021 Gossip and competitive altruism support cooperation in a Public Good game. Phil. Trans. R. Soc. B 376, 20200303.

Giardini, F. and Vilone, D. 2016. Evolution of gossip-based indirect reciprocity on a bipartite network. Sci. Rep. 6, 37931 .

Gintis, H., 2000. Strong reciprocity and human sociality. Journal of theoretical biology, 206(2), pp.169-179.

Gintis, H., Bowles, S., Boyd, R.T. and Fehr, E. eds., 2005. Moral sentiments and material interests: The foundations of cooperation in economic life (Vol. 6). MIT press.

Gneezy, U., 2005. Deception: The role of consequences. Amer. Econ. Rev. 95, 384–394.

Gneezy, U., Meier, S. and Rey-Biel, P. 2011. When and why incentives (don't) work to modify behavior. J. Econ. Perspect. 25, 191–210 .

Goeree, Jacob K et al. 2010. The 1/d law of giving". In: American Economic Journal: Microeconomics 2.1, pp. 183–203.

Grafen, A., 1990. Biological signals as handicaps. Journal of theoretical biology, 144(4), pp.517-546.

Griskevicius, V., Tybur, J. M., and Van den Bergh, B. 2010. Going green to be seen: status, reputation, and conspicuous conservation. Journal of personality and social psychology, 98(3), 392.

Guala, F., 2016. Understanding institutions. In Understanding Institutions. Princeton University Press.

Guimarães, M., MunguÍa-Steyer, R., Doherty Jr, P. F., and Sawaya, R. J. 2017. No survival costs for sexually selected traits in a polygynous non-territorial lizard. Biological Journal of the Linnean Society, 122(3), 614-626.

Haldane, John BS (1955). Population genetics. In: New biology 18.1, pp. 34–51.

Hamilton, William D (1963). The evolution of altruistic behavior. In: The American Naturalist 97.896, pp. 354–356.

Hamilton, William D (1964). The genetical evolution of social behaviour. II. In: Journal of theoretical biology 7.1, pp. 17–52.

Hardin, G., 1968. The tragedy of the commons: the population problem has no technical solution; it requires a fundamental extension in morality. science, 162(3859), pp.1243-1248.

Harkins SG, Petty RE. 1987 Information utility and the multiple source effect. J. Pers. Soc. Psychol. 52, 260.

Haselager, Gerbert JT et al. 1998. Similarities between friends and nonfriends in middle childhood". In: Child Development 69.4, pp. 1198–1208.

Hauk, Esther and Rosemarie Nagel 2001. Choice of partners in multiple two-person prisoner's dilemma games: An experimental study". In: Journal of conflict resolution 45.6, pp. 770–793.

Hawkes, K., and Bliege Bird, R. 2002. Showing off, handicap signalling, and the evolution of men's work. Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews, 11(2), 58-67.

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

Hemelrijk CK, Ek A. 1991 Reciprocity and interchange of grooming and 'support' in captive chimpanzees. Anim. Behav. 41, 923-935.

Henrich, J., Chudek, M. and Boyd, R. 2015. The Big Man Mechanism: how prestige fosters cooperation and creates prosocial leaders. Philos.Trans. R. Soc. B Biol. Sci. 370, 20150013

Henrich, Joseph et al. 2006. Costly punishment across human societies". In: Science 312.5781, pp. 1767–1770.

Herman-Stahl, Mindy and Anne C Petersen 1996. The protective role of coping and social resources for depressive symptoms among young adolescents". In: Journal of youth and Adolescence 25.6, pp. 733–753.

Herrmann, Benedikt, Christian Thoni, and Simon Gachter 2008. Antisocial punishment across societies". In: Science 319.5868, pp. 1362–1367.

Herrmann, E., Engelmann, J. M. and Tomasello, M. 2019. Children engage in competitive altruism. J. Exp. Child Psychol. 179, 176–189 .

Hess NH, Hagen EH. 2021 Competitive gossip: the impact of domain, resource value, resource scarcity and coalitions. Phil. Trans. R. Soc. B 376, 20200305.

Hess, N. H. and Hagen, E. H. 2006. Psychological adaptations for assessing gossip veracity. Hum. Nat. 17, 337–354 .

Hess, N. H., and Hagen, E. H. 2006. Sex differences in indirect aggression: Psychological evidence from young adults. Evolution and Human Behavior, 27(3), 231-245.

Higham, J. P. 2014. How does honest costly signalling work? Behavioral Ecology, 25(1), 8-11.

Hooker, Christine I et al. (2008). Mentalizing about emotion and its relationship to empathy. In: Social cognitive and affective neuroscience 3.3, pp. 204–217.

Horn, Dániel, Hubert János Kiss, and Tünde Lénárd 2022. Gender differences in preferences of adolescents: Evidence from a large-scale classroom experiment". In: Journal of Economic Behavior and Organization 194, pp. 478–522.

Hurd, P. L. 1995. Communication in discrete action-response games. Journal of Theoretical Biology, 174(2), 217-222.

Janky, B., 2022. Erkölcs, racionalitás, közösség: A morális dilemmák egy szociológiai elmélete, Gondolat Kiadó

Johnsen, A. A. and Kvaløy, O. 2016. Does strategic kindness crowd out prosocial behavior?. J. Econ. Behav. Organ. 132, 1–11 .

Kahneman, Daniel, Jack L Knetsch, and Richard H Thaler 1986. Fairness and the assumptions of economics". In: Journal of business, S285–S300.

Kandori, M. 1992. Social norms and community enforcement. Rev. Econ. Stud. 59, 63–80 .

Kantner, J., and Vaughn, K. J. 2012. Pilgrimage as costly signal: Religiously motivated cooperation in Chaco and Nasca. Journal of Anthropological Archaeology, 31(1), 66-82.

Kashy, D. A. and DePaulo, B. M. Who lies? Journal of Personality and Social Psychology, 70, 1037 1996.

Kauser S, Shaw V. 2004 The influence of behavioural and organisational characteristics on the success of international strategic alliances. Int. Mark. Rev. 21, 17-52.

Kirkpatrick, M. 1986. The handicap mechanism of sexual selection does not work. The American Naturalist. 127, 222–240.

Knifsend, Casey A et al. 2018. Friends in activities, school-related affect, and academic outcomes in diverse middle schools”. In: Journal of youth and adolescence 47.6, pp. 1208–1220.

Kollock, P., 1998. Social dilemmas: The anatomy of cooperation. Annual review of sociology, pp.183-214.

Kotiaho, J. S. 2001. Costs of sexual traits: a mismatch between theoretical considerations and empirical evidence. Biological Reviews, 76(3), 365-376.

Kovářík, J., Brañas-Garza, P., Cobo-Reyes, R., Espinosa, M.P., Jiménez, N. and Ponti, G., 2012. Prosocial norms and degree heterogeneity in social networks”. In: Physica A: Statistical Mechanics and its Applications 391.3, pp. 849–853.

Kuttler AF, Parker JG, La Greca AM. 2002 Developmental and gender differences in preadolescents' judgments of the veracity of gossip. Merrill-Palmer Q. 48, 105-132.

Lachmann, M., Számadó, S., and Bergstrom, C. T. 2001. Cost and conflict in animal signals and human language. Proceedings of the National Academy of Sciences, 98(23), 13189-13194.

Laidre ME, Lamb A, Shultz S, Olsen M. 2013 Making sense of information in noisy networks: human communication, gossip, and distortion. J. Theor. Biol. 317, 152-160.

Lave, L.B., 1962. An empirical approach to the prisoners' dilemma game. The Quarterly Journal of Economics, pp.424-436.

Ledyard, John O and Thomas R Palfrey 1994. Voting and lottery drafts as efficient public goods mechanisms”. In: The Review of Economic Studies 61.2, pp. 327–355.

Leider, Stephen et al. 2009. Directed altruism and enforced reciprocity in social networks”. In: The Quarterly Journal of Economics 124.4, pp. 1815–1851.

Levine, E. E. and Schweitzer, M. E. 2015. Prosocial lies: when deception breeds trust. Organ. Behav. Hum. Decis. Process. 126, 88–106.

Lewis, D. K. (1969). Convention. Cambridge, MA: Harvard University Press.

Logis, Handrea A et al. 2013. Popularity as an organizing factor of preadolescent friendship networks: Beyond prosocial and aggressive behavior". In: Journal of Research on Adolescence 23.3, pp. 413–423.

Lomi, Alessandro et al. 2011. Why are some more peer than others? Evidence from a longitudinal study of social networks and individual academic performance". In: Social science research 40.6, pp. 1506–1520.

Luce, R. D. and Raiffa, H. 1957. Games and Decisions Wiley, New York,.

Lyle III, H. F., Smith, E. A., and Sullivan, R. J. 2009. Blood donations as costly signals of donor quality. Journal of Evolutionary Psychology, 7(4), 263-286.

Lyle, H. F. and Smith, E. A. 2014. The reputational and social network benefits of prosociality in an Andean community. Proc. Natl. Acad.Sci. U.S.A. 111, 4820–4825 .

Macfarlan, S. J., Quinlan, R. and Remiker, M. 2013. Cooperative behaviour and prosocial reputation dynamics in a Dominican village.Proc. R. Soc. B Biol. Sci. 280, 20130557 .

Macfarlan, Shane J and Henry F Lyle 2015. Multiple reputation domains and cooperative behaviour in two Latin American communities". In: Philosophical Transactions of the Royal Society B: Biological Sciences 370.1683, p. 20150009.

Maynard Smith, J. 1976. Sexual Selection and the Handicap Principle. Journal of Theoretical Biology. 57, 239–242.

Maynard Smith, J. M., and Harper, D. 2003. Animal signals. Oxford University Press.

McAndrew, F. T. 2014. The, “sword of a woman”: gossip and female aggression. Aggress. Violent Behav. 19, 196–199 .

McAndrew, F. T. 2018. Costly signalling theory. Todd K. Shackelford and Viviana A. Weekes-Shackelford, Encyclopedia of Evolutionary Psychological Science.

McAndrew, F.T., 2002. New evolutionary perspectives on altruism: Multilevel-selection and costly-signaling theories. Current Directions in Psychological Science, 11(2), pp.79-82.

McCarty, J. P. 1996. The energetic cost of begging in nestling passerines. The AUK, 113(1), 178-188.

McCullough, E. L., and Emlen, D. J. 2013. Evaluating the costs of a sexually selected weapon: big horns at a small price. Animal Behaviour, 86(5), 977-985.

McPherson, Miller, Lynn Smith-Lovin, and James M Cook 2001. Birds of a feather: Homophily in social networks”. In: Annual review of sociology 27.1, pp. 415–444.

Melamed, David, Ashley Harrell, and Brent Simpson 2018. Cooperation, clustering, and assortative mixing in dynamic networks”. In: Proceedings of the National Academy of Sciences 115.5, pp. 951–956.

Melamed, David, Brent Simpson, and Ashley Harrell 2017. Prosocial orientation alters network dynamics and fosters cooperation”. In: Scientific reports 7.1, pp. 1–6.

Melis AP, Semmann D. 2010 How is human cooperation different?Phil. Trans. R. Soc. B 365, 2663-2674.

Mercier H. 2012 The social functions of explicit coherence evaluation. Mind Soc. 11, 81-92.

Merton, R.K. and Merton, R.C., 1968. Social theory and social structure. Simon and Schuster.

Milinski M. 2016 Reputation, a universal currency for human social interactions. Phil. Trans. R. Soc. B 371, 20150100.

Milinski, M., Semmann, D. and Krambeck, H. J. 2002. Reputation helps solve the ’tragedy of the commons’. Nature 415, 424–426 .

Milinski, M., Semmann, D. and Krambeck, H., 2002. Donors to charity gain in both indirect reciprocity and political reputation. Proceedings of the Royal Society of London. Series B: Biological Sciences, 269(1494), pp.881-883.

Miller, D.T. and Ratner, R.K., 1998. The disparity between the actual and assumed power of self-interest. Journal of personality and social psychology, 74(1), p.53.

Mills CM, Grant MG. 2009 Biased decision-making: developing an understanding of how positive and negative relationships may skew judgments. Dev. Sci. 12, 784-797.

Molano, Andres et al. 2013. Selection and socialization of aggressive and prosocial behavior: The moderating role of social-cognitive processes”. In: Journal of Research on Adolescence 23.3, pp. 424–436.

Molho C, Wu J. 2021 Direct punishment and indirect reputation-based tactics to intervene against offences. Phil. Trans. R. Soc. B 376, 20200289.

Moreno-Rueda, G. 2007. Is there empirical evidence for the cost of begging?. Journal of Ethology, 25(3), 215-222.

Mulder, L. B., van Dijk, E., De Cremer, D. and Wilke, H. A. M. 2006. Undermining trust and cooperation: the paradox of sanctioning systems in social dilemmas. J. Exp. Soc. Psychol. 42, 147–162 .

Nell, V. 2002. Why young men drive dangerously: Implications for injury prevention. Current Directions in Psychological Science, 11(2), 75-79.

Newcomb, A.F., Bukowski, W.M. and Pattee, L., 1993. Children's peer relations: a meta-analytic review of popular, rejected, neglected, controversial, and average sociometric status. Psychological bulletin, 113(1), p.99.

Newcomb, Andrew F and Catherine L Bagwell 1995. Children's friendship relations: A meta-analytic review." In: Psychological bulletin 117.2, p. 306.

Nikiforakis, Nikos 2008. Punishment and counter-punishment in public good games: Can we really govern ourselves?" In: Journal of Public Economics 92.1-2, pp. 91–112.

Nowak, M. A. 2006. Five rules for the evolution of cooperation. Science 314, 1560–1563.

Nowak, M. A. and Sigmund, K. 1998. Evolution of indirect reciprocity by image scoring. Nature 393, 573–577.

Nowak, M. A. and Sigmund, K. 2005. Evolution of indirect reciprocity. Nature 437, 1291–1298.

Nowak, Martin A (2006). Five rules for the evolution of cooperation. In: science 314.5805, pp. 1560– 1563.

OECD, 2019. Evaluating Latvia's Active Labour Market Policies, Connecting People with Jobs, OECD Publishing, Paris.

Ohtsuki, H. and Iwasa, Y. 2004. How should we define goodness?—reputation dynamics in indirect reciprocity. J. Theor. Biol. 231, 107–120.

Ohtsuki, H. and Iwasa, Y. 2006. The leading eight: social norms that can maintain cooperation by indirect reciprocity. J. Theor. Biol. 239,435–444.

Ohtsuki, H. and Iwasa, Y. 2007. Global analyses of evolutionary dynamics and exhaustive search for social norms that maintain cooperation by reputation. J. Theor. Biol. 244, 518–531 .

Ohtsuki, H., Iwasa, Y. and Nowak, M. A. 2015. Reputation effects in public and private interactions. PLoS Comput. Biol. https://doi.org/10.1371/journal.pcbi.1004527 .

Ohtsuki, H., Iwasa, Y. and Nowak, M.A., 2009. Indirect reciprocity provides only a narrow margin of efficiency for costly punishment. Nature, 457(7225), pp.79-82.

Olson, M. 1965. The Logic of Collective Action Harvard University Press, Cambridge.

Ostrom E, Burger J, Field CB, Norgaard RB, Policansky D. 1999. Revisiting the commons: local lessons, global challenges. Science 284, 278-282.

Paine, R. 1967. What is gossip about? An alternative hypothesis. Man 2, 278–285.

Pál, J., Stadtfeld, C., Grow, A. and Takács, K., 2016. Status perceptions matter: Understanding disliking among adolescents. Journal of Research on Adolescence, 26(4), pp.805-818.

Panchanathan K, Boyd R. 2004 Indirect reciprocity can stabilize cooperation without the second-order free rider problem. Nature 432, 499-502.

Papadopoulos Y. 2003 Cooperative forms of governance: problems of democratic accountability in complex environments. Eur. J. Polit. Res. 42, 473-501.

Pasquini ES, Corriveau KH, Koenig M, Harris PL. 2007 Preschoolers monitor the relative accuracy of informants. Dev. Psychol. 43, 1216.

Penn, D.J. and Számadó, S., 2020. The Handicap Principle: how an erroneous hypothesis became a scientific principle. Biological Reviews, 95(1), pp.267-290.

Perc, Matjaž and Attila Szolnoki 2010. Coevolutionary games-a mini review. In: BioSystems 99.2, pp. 109–125.

Peters K, Fonseca MA. 2020 Truth, lies, and gossip. Psychol. Sci. 31, 702-714.

Piazza J, Bering JM. 2008 Concerns about reputation via gossip promote generous allocations in an economic game. Evol. Hum. Behav. 29, 172-178.

Pletzer, Jan Luca et al. 2018. Social value orientation, expectations, and cooperation in social dilemmas: A meta–analysis". In: European Journal of Personality 32.1, pp. 62–83.

Podder S, Righi S, Pancotto F. 2021 Reputation and punishment sustain cooperation in the optional public goods game. Phil. Trans. R. Soc. B 376, 20200293.

Polnaszek, T. J., and Stephens, D. W. 2014. Why not lie? Costs enforce honesty in an experimental signalling game. Proceedings of the Royal Society B: Biological Sciences, 281(1774), 20132457.

Posner, E.A., 2000. Law and social norms, Harvard University Press

Przepiorka W, Diekmann A. 2021 Parochial cooperation and the emergence of signalling norms. Phil. Trans. R. Soc. B 376, 20200294.

Raihani NJ, Bshary R. 2015 Why humans might help strangers. Front. Behav. Neurosci. 9, 39.

Raihani, N. J. and Bshary, R. 2015. The reputation of punishers. Trends Ecol. Evol. 30, 98–103 .

Raihani, N.J. and Smith, S., 2015. Competitive helping in online giving. Current Biology, 25(9), pp.1183-1186.

Rand, D. G. and Nowak, M. A. 2013. Human cooperation. Trends Cogn. Sci. 17, 413–425 .

Rand, David G 2016. Cooperation, fast and slow: Meta-analytic evidence for a theory of social heuristics and self-interested deliberation". In: Psychological science 27.9, pp. 1192–1206.

Rand, David G, Alexander Peysakhovich, et al. 2014. Social heuristics shape intuitive cooperation”. In: Nature communications 5.1, pp. 1–12.

Rand, David G, Joshua D Greene, and Martin A Nowak 2012. Spontaneous giving and calculated greed”. In: Nature 489.7416, pp. 427–430.

Rand, David G, Samuel Arbesman, and Nicholas A Christakis 2011. Dynamic social networks promote cooperation in experiments with humans”. In: Proceedings of the National Academy of Sciences 108.48, pp. 19193–19198.

Rapoport A. 1967 A note on the ’index of cooperation’ for Prisoner's Dilemma. J. Confl. Resolut. 11, 100-103.

Rapoport, A. and Chammah, A.M., 1970. Prisoner's dilemma: A study in conflict and cooperation (Vol. 165). University of Michigan press.

Righi, S. and Takács, K. 2018. Social closure and the evolution of cooperation via indirect reciprocity. Sci. Rep. 8, 1–9 .

Riley, J. G. 1979. Informational equilibrium. Econometrica, 331-359.

Ripley, Ruth M et al. 2011. Manual for RSIENA. In: University of Oxford, Department of Statistics, Nuffield College 1, p. 2011.

Roberts AI, Roberts SG. 2017 Convergence and divergence in gesture repertoires as an adaptive mechanism for social bonding in primates. R. Soc. Open Sci. 4, 170181.

Roberts G, Raihani N, Bshary R, Manrique HM, Farina A, Samu F, Barclay P. 2021 The benefits of being seen to help others: indirect reciprocity and reputation-based partner choice. Phil. Trans. R. Soc. B 376, 20200290.

Roberts G. 2008 Evolution of direct and indirect reciprocity. Proc. R. Soc. B 275, 173-179.

Roberts, G. 1998. Competitive altruism: from reciprocity to the handicap principle. Proceedings of the Royal Society of London. Series B: Biological Sciences, 265(1394), 427-431.

Romano A, Giardini F, Columbus S, de Kwaadsteniet EW, Kisfalusi D, Triki Z, Snijders C, Hagel K. 2021 Reputation and socio-ecology in humans. Phil. Trans. R. Soc. B 376, 20200295.

Rubin, H., Bruner, J. P., O'Connor, C., and Huttegger, S. 2020) "Communication without Common Interest: A Signalling Experiment." Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences. 83 October): 101295. https://doi.org/10.1016/j.shpsc.2020.101295.

Salamon, S. D., and Deutsch, Y. 2006. OCB as a handicap: An evolutionary psychological perspective. Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior, 27(2), 185-199.

Samu F, Számadó S, Takács K. 2020 Scarce and directly beneficial reputations support cooperation. Sci. Rep. 10, 1-2.

Sanchez-Pages, S., and M. Vorsatz. 2007) "An Experimental Study of Truth-Telling in a Sender-Receiver Game." Games and Economic Behavior. 61 2007): 86–112. https://doi.org/10.1016/j.geb.2006.10.014.

Santos, Francisco C, Jorge M Pacheco, and Tom Lenaerts 2006. Evolutionary dynamics of social dilemmas in structured heterogeneous populations". In: Proceedings of the National Academy of Sciences 103.9, pp. 3490–3494.

Santos, Francisco C, Marta D Santos, and Jorge M Pacheco 2008. Social diversity promotes the emergence of cooperation in public goods games". In: Nature 454.7201, pp. 213–216.

Searcy, W. A. and Nowicki, S. 2005. The Evolution of Animal Communication. Reliability and Deception in Signalling Systems. Princeton University Press, Princeton.

Seinen, I. and Schram, 2006. A. Social status and group norms: indirect reciprocity in a repeated helping experiment. Eur. Econ. Rev. 50,581–602 .

Semmann, D., Krambeck, H.-J. and Milinski, M. 2005. Reputation is valuable within and outside one's own social group. Behav. Ecol.Sociobiol. 57, 611–616 .

Shalvi, S. and De Dreu, C. K. 2014. Oxytocin promotes group-serving dishonesty. Proc. Natl. Acad. Sci. 111, 5503–5507 .

Shimada M, Sueur C. 2018 Social play among juvenile wild Japanese macaques (Macaca fuscata) strengthens their social bonds. Am. J. Primatol. 80, e22728.

Shin, Huiyoung, Allison M Ryan, and Elizabeth North 2019. Friendship processes around prosocial and aggressive behaviors: The role of teacher–student relatedness and differences between elementary-school and middle-school classrooms". In: Merrill-Palmer Quarterly 65.2, pp. 232–263.

Simpson, B. and Willer, R., 2015. Beyond altruism: Sociological foundations of cooperation and prosocial behavior. Annual Review of Sociology, 41(1), pp.43-63.

Singer, T. and Fehr, E., 2005. The neuroeconomics of mind reading and empathy. American Economic Review, 95(2), pp.340-345.

Singer, Tania and Ernst Fehr (2005). The neuroeconomics of mind reading and empathy. In:American Economic Review 95.2, pp. 340–345.

Sivanathana N. and Pettit N.C. 2010 Protecting the self through consumption: Status goods as affirmational commodities. Journal of Experimental Social Psychology, Volume 46, Issue 3, 564-570.

Skyrms, B., 2004. The stag hunt and the evolution of social structure. Cambridge University Press.

Smith EA. 2010 Communication and collective action: language and the evolution of human cooperation. Evol. Hum. Behav. 31, 231-245.

Smith ER. 2014 Evil acts and malicious gossip: a multiagent model of the effects of gossip in socially distributed person perception. Pers. Soc. Psychol. Rev. 18, 311-325.

Smith, E. A. and Bird, R. L. B. 2000. Turtle hunting and tombstone opening: public generosity as costly signaling. Evol. Hum. Behav. 21,245–261 .

Smith, J (1964). Group selection and kin selection. In: Nature 201.4924, pp. 1145–1147.

Snellman, J.E., Iñiguez, G., Kertész, J., Barrio, R.A. and Kaski, K.K., 2019. Status maximization as a source of fairness in a networked dictator game. Journal of Complex Networks, 7(2), pp.281-305.

Snijders, Tom AB, Gerhard G Van de Bunt, and Christian EG Steglich 2010. Introduction to stochastic actor-based models for network dynamics". In: Social networks 32.1, pp. 44–60.

Sommerfeld, R. D., Krambeck, H.-J. and Milinski, M. 2008. Multiple gossip statements and their effect on reputation and trustworthiness.Proc. R. Soc. B Biol. Sci. 275, 2529–2536 .

Sommerfeld, R. D., Krambeck, H.-J., Semmann, D. and Milinski, M. 2007. Gossip as an alternative for direct observation in games of indirect reciprocity. Proc. Natl. Acad. Sci. U.S.A. 104, 17435–17440 .

Sosis, R. 2003. Why aren't we all hutterites?. Human Nature, 14(2), 91-127.

Sosis, R., and Bressler, E. R. 2003. Cooperation and commune longevity: A test of the costly signalling theory of religion. Cross-Cultural Research, 37(2), 211-239.

Spence, M. 1973. Job market signalling. In Uncertainty in Economics pp. 281-306.

Academic Press.

Stadtfeld, Christoph, Károly Takács, and András Vörös 2020. The emergence and stability of groups in social networks”. In: Social Networks 60, pp. 129–145.

Stibbard-Hawkes, D. N. 2019. Costly signalling and the handicap principle in hunter-gatherer research: A critical review. Evolutionary Anthropology: Issues, News, and Reviews, 28(3), 144-157.

Sugden, R. et al. The Economics of Rights, Co-operation and Welfare (Springer, Berlin, 2004).

Swakman V, Molleman L, Ule A, Egas M. 2016 Reputation-based cooperation: empirical evidence for behavioral strategies. Evol. Hum. Behav. 37, 230-235.

Sylwester, K. and Roberts, G. 2013. Reputation-based partner choice is an effective alternative to indirect reciprocity in solving social dilemmas. Evol. Hum. Behav. 34, 201–206 .

Sylwester, K. and Roberts, G., 2010. Cooperators benefit through reputation-based partner choice in economic games. Biology letters, 6(5), pp.659-662.

Sylwester, Karolina, Benedikt Herrmann, and Joanna J Bryson 2013. Homo homini lupus? Explaining antisocial punishment.” In: Journal of Neuroscience, Psychology, and Economics 6.3,p. 167.

Számadó, S. 1999. The validity of the handicap principle in discrete action–response games. Journal of Theoretical Biology, 198(4), 593-602.

Számadó, S. 2003. Threat displays are not handicaps. Journal of Theoretical Biology, 221(3), 327-348.

Számadó, S. 2011. The cost of honesty and the fallacy of the handicap principle. Animal

Behaviour, 81(1), 3-10.

Számadó, S., and Penn, D. J. 2015. Why does costly signalling evolve? Challenges with testing the handicap hypothesis. Animal Behaviour, 110, e9.

Számadó, S., and Penn, D. J. 2018. Does the handicap principle explain the evolution of dimorphic ornaments? Animal Behaviour, 138, e7.

Számadó, S., Czégel, D., and Zachar, I. 2019. One problem, too many solutions: How costly is honest signalling of need?. PloS one, 14(1), e0208443.

Számadó, S., Szalai, F. and Scheuring, I. 2016. Deception undermines the stability of cooperation in games of indirect reciprocity. PLoSONE 11, e0147623 .

Takács K, Gross J, Testori M, Letina S, Kenny AR, Power EA, Wittek RPM. 2021 Networks of reliable reputations and cooperation: a review. Phil. Trans. R. Soc. B 376, 20200297.

Takács, Károly, Béla Janky, and Andreas Flache 2008. Collective action and network change”. In: Social Networks 30.3, pp. 177–189.

Thavarajah, N. K., Tickle, P. G., Nudds, R. L., and Codd, J. R. 2016. The peacock train does not handicap cursorial locomotor performance. Scientific Reports, 6(1), 1-6.

Thielmann, Isabel, Giuliana Spadaro, and Daniel Balliet 2020. Personality and prosocial behavior: A theoretical framework and meta-analysis.” In: Psychological Bulletin 146.1, p. 30.

Torsvik G, Molander A, Tjøtta S, Kobbeltvedt T. 2011 Anticipated discussion and cooperation in a social dilemma. Ration. Soc. 23, 199-216.

Uehara, E., 1990. Dual exchange theory, social networks, and informal social support. American Journal of Sociology, 96(3), pp.521-557.

Ueno, Koji 2005. The effects of friendship networks on adolescent depressive symptoms”. In: Social Science Research 34.3, pp. 484–510.

Van de Bunt GG, Wittek RP, de Klepper MC. 2005 The evolution of intra-organizational trust networks: the case of a German paper factory: an empirical test of six trust mechanisms. Int. Sociol. 20, 339-369.

Van Lange, P. A. M., Balliet, D., Parks, C. D. and van Vugt, M. 2014. Social Dilemmas: Understanding Human Cooperation, Oxford University Press, New York.

Veblen, Thorstein 1899. The Theory of the Leisure Class: An Economic Study of Institutions. New York: The Macmillan Company.

Viechtbauer, Wolfgang 2010. Conducting meta-analyses in R with the metafor package”. In: Journal of statistical software 36.3, pp. 1–48.

von Rueden, C.R., Redhead, D., O'Gorman, R., Kaplan, H. and Gurven, M., 2019. The dynamics of men's cooperation and social status in a small-scale society. Proceedings of the Royal Society B, 286(1908), p.20191367.

Vörös András, Per Block, and Zsófia Boda 2019. Limits to inferring status from friendship relations”. In: Social Networks 59, pp. 77–97.

Vörös, András and Tom AB Snijders 2017. Cluster analysis of multiplex networks: Defining composite network measures”. In: Social Networks 49, pp. 93–112.

Wang, Jing, Siddharth Suri, and Duncan J Watts 2012. Cooperation and assortativity with dynamic partner updating”. In: Proceedings of the National Academy of Sciences 109.36, pp. 14363–14368.

Wang, M. Y., Vasas, V., Chittka, L., and Yen, S. H. 2017. Sheep in wolf's clothing: multicomponent traits enhance the success of mimicry in spider-mimicking moths. Animal Behaviour, 127, 219-224.

Weber M. 2020[1920] A protestáns etika és a kapitalizmus szelleme, L' HARMATTAN KFT.

Webster, M. and Sell, J. eds., 2014. Laboratory experiments in the social sciences. Elsevier.

Wedekind, C. and Milinski, M. 2000. Cooperation through image scoring in humans. Science (New York, N.Y.) 288, 850–852 .

Willer, Robb 2009. Groups reward individual sacrifice: The status solution to the collective action problem". In: American Sociological Review 74.1, pp. 23–43.

Wilson DS, Wilczynski C, Wells A, Weiser L. 2000 Gossip and other aspects of language as group-level adaptations. In The evolution of cognition (eds CM Heyes, L Huber), pp. 347-365. Cambridge, MA: MIT Press.

Wittig RM, Crockford C, Deschner T, Langergraber KE, Ziegler TE, Zuberbühler K. 2014 Food sharing is linked to urinary oxytocin levels and bonding in related and unrelated wild chimpanzees. Proc. R. Soc. B 281, 20133096

Wu J, Balliet D, Peperkoorn LS, Romano A, Van Lange PA. 2020 Cooperation in groups of different sizes: the effects of punishment and reputation-based partner choice. Front. Psychol. 10, 2956.

Wu J, Balliet D, Van Lange PA. 2016 Reputation management: why and how gossip enhances generosity. Evol. Hum. Behav. 37, 193-201.

Wu J, Balliet D, Van Lange PA. 2016 Reputation, gossip, and human cooperation. Soc. Pers. Psychol. Compass 10, 350-364.

Yamagishi, T., Mifune, N., Li, Y., Shinada, M., Hashimoto, H., Horita, Y., Miura, A., Inukai, K., Tanida, S., Kiyonari, T. and Takagishi, H., 2013. Is behavioral pro-sociality game-specific? Pro-social preference and expectations of pro-sociality. Organizational Behavior and Human Decision Processes, 120(2), pp.260-271.

Yamagishi, Toshio et al. (2013). Is behavioral pro-sociality game-specific? Pro-social preference and expectations of pro-sociality. In: Organizational Behavior and Human Decision Processes 120.2, pp. 260–271.

Yoeli, E., Hoffman, M., Rand, D. G. and Nowak, M. 2013. A. Powering up with indirect reciprocity in a large-scale field experiment. Proc.Natl. Acad. Sci. 110, 10424–10429 .

Zahavi, A. 1975. Mate selection—a selection for a handicap. Journal of Theoretical Biology, 53(1), 205-214.

Zahavi, A. and Zahavi, A., 1999. The handicap principle: A missing piece of Darwin's puzzle. Oxford University Press.

Zahavi, A., and Zahavi, A. 1997. The handicap principle: A missing piece of Darwin's puzzle. Oxford University Press.

Zand DE. 1972 Trust and managerial problem solving. Adm. Sci. Q. 17, 229-239.

Zizzo, D.J., 2004. Inequality and procedural fairness in a money-burning and stealing experiment. In Inequality, welfare and income distribution: Experimental approaches. Emerald Group Publishing Limited.

Zizzo, Daniel John 2003. Money burning and rank egalitarianism with random dictators”. In: Economics Letters 81.2, pp. 263–266.

Zollman, K. J., Bergstrom, C. T., and Huttegger, S. M. 2013. Between cheap and costly signals: the evolution of partially honest communication. Proceedings of the Royal Society B: Biological Sciences, 280(1750), 20121878.

8. LIST OF PUBLICATIONS

8.1. Peer-reviewed journal articles

Samu, F. and Takács, K., 2021. Evaluating mechanisms that could support credible reputations and cooperation: cross-checking and social bonding. Philosophical Transactions of the Royal Society B, 376(1838), p.20200302.

Samu, F., Számadó, S. and Takács, K., 2020. Scarce and directly beneficial reputations support cooperation. Scientific reports, 10(1), pp.1-12.

Számadó, S., Samu, F., & Takács, K. 2022. Condition-dependent trade-offs maintain honest signalling. Royal Society Open Science, 9(10), 220335.

Roberts, G., Raihani, N., Bshary, R., Manrique, H.M., Farina, A., Samu, F. and Barclay, P., 2021. The benefits of being seen to help others: indirect reciprocity and reputation-based partner choice. Philosophical Transactions of the Royal Society B, 376(1838), p.20200290.

Manrique, H.M., Zeidler, H., Roberts, G., Barclay, P., Walker, M., Samu, F., Fariña, A., Bshary, R. and Raihani, N., 2021. The psychological foundations of reputation-based cooperation. Philosophical Transactions of the Royal Society B, 376(1838), p.20200287.

Herke, B., and Samu, F. (2019). Wim van Oorschot, Femke Roosma, Bart Meuleman and Tim Reeskens (eds)(2017), The Social Legitimacy of Targeted Welfare: Attitudes to Welfare Deservingness, Cheltenham: Edward Elgar. Journal of Social Policy, 48(3), 648-650.

8.2. International conference presentations

Flóra Samu, Szabolcs Számadó and Károly Takács (2018). Scarce and directly beneficial reputations support cooperation. EHBEA, Hungary, Pécs

Flóra Samu, Dorottya Kisfalusi and Károly Takács (2018) Well-Being and the Evolution of Positive Relations in School, NETGLOW, Saint Petersburg, Russia

Flóra Samu (2019) The Dynamic of Gossip: an Empirical Study How Interpersonal Communication About Others Changes Opinion-Based Support EUSN, Zürich, Switzerland

Flóra Samu (2020). Social Preferences and the Evolution of Friendship and Negative Relations in Primary Schools, SUNBELT, online (Paris, France)

8.3. Others

Samu, F., Scharle, Á. and Csillag, M. 2019. Female employment in post-socialist EU member states In: Fazekas, K; Szabó-Morvai, Á (eds.) The Hungarian Labour Market, 2018 Budapest, Hungary, Institute of Economics, Centre for Economic and Regional Studies, Hungarian Academy of Sciences, pp. 71-75., 5 p.

Herke, B., and Samu, F. 2019. Wim van Oorschot, Femke Roosma, Bart Meuleman and Tim Reeskens (eds)(2017), The Social Legitimacy of Targeted Welfare: Attitudes to Welfare Deservingness, Cheltenham: Edward Elgar. Journal of Social Policy, 48(3), 648-650. (Book review)

Scharle, Á., Váradi, B., and Samu, F. 2015. Policy convergence across welfare regimes: the case of disability policies (No. 76). WWWforEurope Working Paper.

Csillag, M., Samu, F., and Scharle, Á. 2014. Exclusion from the labour market in post-socialist EU member states. Grincoh Working Paper.