The language of cooperation: reputation and honest signalling

Számadó, S.[1,2], Balliet, D.[3], Giardini, F.[4], Power, E. A.[5], Takács, K.[6,2]

1 Department of Sociology and Communication, BUTE, 2 CSS-RECENS, Centre for Social Sciences, 3 Vrije Universiteit Amsterdam, 4 University of Groningen, 5 London School of Economics & Political Science, 6 The Institute for Analytical Sociology, Linköping University

24 August 2021

Abstract

Large-scale non-kin cooperation is a unique ingredient of human success. This type of cooperation is challenging to explain in a world of self-interested individuals. There is overwhelming empirical evidence from different disciplines that reputation and gossip promotes cooperation in humans in different contexts. Despite decades of research important details of reputation systems are still unclear. Our goal with this theme issue is to promote and inter-disciplinary approach that allows us to explore and to understand the evolution and the maintenance of reputation system with special emphasis on gossip and honest signalling. The theme issue is organised along four main questions: What are the necessary conditions for reputation-based systems? What is the content and context of reputation systems? How can reputations promote cooperation? What is the role of gossip in maintaining reputation systems and thus cooperation?

This article is part of the theme issue 'The language of cooperation: reputation and honest signalling'

1. Introduction

Cooperation among unrelated members is one of the most prominent features of human societies that makes life together highly successful (Maynard Smith, Szathmáry, 1995; Noë, Hammerstein, 1995; Granovetter, 1978; Kollock, 1998). Assigning reputation to others, sharing reputations of others, and conditioning behaviour on reputation are among the main driving forces of such cooperation (Nowak, Sigmund, 1998; Milinski et al., 2002; Panchanathan, Boyd, 2004; Giardini, Wittek, 2019a). While the effect and role of reputations on cooperation is widely studied (Nowak, Sigmund, 2005; Tennie, Frith, Frith, 2010; Simpson, Willer, 2015; Barclay, 2012; Wu, Balliet, & Van Lange, 2016a; Milinski, 2016; Bliege Bird, Ready, Power, 2018; Marsh, 2018; Giardini, Wittek, 2019b), challenging questions are left unanswered about the evolution and conditions of reputation systems, the context and content of reputation, and the operation of reputation-based systems. Signaling and communication are essential for reputations being shared, and among humans, third-party sharing of reputations through gossip can amplify the potential of reputations to sustain cooperation. Signals and gossip, however, are not necessarily honest and the informal mechanisms that can safeguard the functioning of reputation systems must be uncovered. In this thematic issue, scholars coming from different backgrounds compare theories, methods and perspectives, and offer new empirical research in order to lay the foundations for an interdisciplinary and systematic research program on reputation, gossip and honesty for sustainable cooperation. The articles in this special issue all draw our attention to these complexities of the workings of reputation systems, asking: (i) What are the necessary conditions for reputation-based systems? (ii) What is the content and context of reputation systems? (iii) How can reputations promote cooperation? And (iv) what is the role of gossip in maintaining reputation systems and thus cooperation?

This theme issue is based on the workshop with the same title held in the Lorentz Center in Leiden (Netherlands, September, 2019). We are very grateful to both the Lorentz Center and the Royal Society for providing the physical and intellectual space for the discussion of this topic.

2. Background

Cooperation is the production of mutual benefits at a cost to the individual. The collective benefits that cooperation can produce help explain the success of societies that establish and sustain cooperation. Yet cooperation can be difficult to achieve, since there can be higher immediate personal costs of cooperation that curb individuals' motivation to cooperate voluntarily. This inherent puzzle of cooperation explains why it is so rare in nature and why it brings about major transitions in evolution once it has been established (Maynard Smith, Szathmáry, 1995). **Large-scale non-kin cooperation** is unique to the human species. While social living species can display an impressive level of coordinated effort, this cooperation typically is fostered by kin selection, where cooperation is promoted by sufficiently high levels of relatedness between individuals (Maynard Smith, Szathmáry, 1995). Human societies are uniquely able to achieve high levels of cooperation, even with low levels of biological relatedness between members of the society. Cooperation in human societies is supported by a complex set of tools, including informal and formal mechanisms. Some mechanisms are able to operate because humans have evolved advanced cognitive and social skills and a complex system of communication that includes language and an opportunity to exchange information about other individuals. A wealth of empirical data from different disciplines demonstrates that this information exchange is crucial in maintaining cooperation (Smith, 2010; Feinberg et al., 2014; Wu et al., 2016a; Shank et al., 2019).

A **reputation** can be seen as an evaluation of an individual based on their past action(s). The importance of such a reputation is that it will determine how others will treat/react to them in future encounters. This reputation can be solely based on **individual** experiences (e.g., based on past interactions and/or observations of the ego), or it can be based on the **aggregated** experiences of a group. The resulting reputations can be **stored privately** (e.g., in the brains of the individuals within the group), allowing different individuals to have different reputational assessments of the same person; or they can be **stored publicly (e.g.**, a public webpage of tax evaders).

A **reputation system** based on aggregated experiences includes (i) a set of decision rules to guide how individuals should be evaluated based on their actions, usually called **social norms**; (ii) a **communication system** that allows for the exchange of experiences and evaluations, and (iii) a set of decision rules that guides the actions towards individuals with different reputations, usually called **behavioural norms** or action rules (Ohtsuki, Iwasa, 2004; 2006; Sigmund, 2012; Okada, 2020; Podder et al., 2021).

3. This issue's contributions

A first and fundamental point at issue is the **evolution of reputation systems**. While some animals have reputation systems based on individual experiences—one famous example is the cleaner wrasse client system (Bshary & Grutter, 2006)—aggregated reputation systems are missing. This begs the question: if aggregated reputation systems are such a powerful tool to promote cooperation, why does it occur only in humans? The answer to this question likely involves the human ability of language. Aggregated reputations are formed through communication, thus the lack of language seems to pose a barrier for other species. Reputation systems require other cognitive skills as well, such as the ability to judge the observed actions of others based on the context of the action. *Manrique et al*. evaluate these cognitive requirements in their review. They argue that a large working memory enables humans to better understand others' mental states, such as through perspective-taking and the attribution of intentions, and which also allows humans to create and follow norms. Therefore, a large working memory may provide the foundation for increasingly complex reputation-based cooperation in humans. They conclude that most non-human species lack a working memory capacity with sufficient computational power to support the formation to form complex systems of reputation-based cooperation.

Inference about the mental states of others plays a key role in our judgements of their actions. For example, evolutionary models show that cooperation can be stabilized by distinguishing between different reasons for the same action, such as selfish defection versus altruistic punishment. However, anthropologists have suggested that some societies seem to disregard mental states when passing judgement. *Barrett and Saxe* argue that this is likely a misapprehension. They propose instead that the extent to which mental states are considered in moral judgment depends on the context and the reasons for judgment. There is no society that would completely disregard mental states, instead it is the importance of mental states in making moral judgement that varies from context to context. It follows that it is important to find out why different societies evaluate actors in the same context differently. They argue that in general there are three important determinants: "Who are the actors?" e.g. individuals vs. corporations; "What are they doing?" e.g. doing inherently dangerous actions vs. non-dangerous actions, and finally "Why is the judgement being made?" e.g. truth making vs. restoring social cohesion.

Models of **Indirect Reciprocity** (IR) attempt to fully specify and analyse a reputation system (IR; Nowak, Sigmund, 1998, 2005; Leimar, & Hammerstein, 2001; Panchanathan & Boyd, 2004; Pacheco

et al, 2006; Ohtsuki & Iwasa, 2004, 2006; Roberts, 2008). Models of indirect reciprocity consider interactions in a large population that prevent the evolution of direct reciprocity ("I scratch your back you scratch mine"). Instead of such direct help, IR models assume a helping system mediated by reputation ('good' vs. 'bad'). Individuals can gain a good reputation by helping others and individuals in good reputation will be helped by other members of the group. While the idea is very simple, details of IR models can be very complex (Santos et al., 2018). *Santos* and colleagues quantify cognitive complexity in IR and evaluate recent advances and potential extensions of indirect reciprocity models. They identify three determinants of cooperation under IR that can affect both the level of cooperation and the level of complexity and information requirements for IR to promote cooperation. These determinants are: (i) costly reputation spread, where individuals have to pay a cost to share information; (ii) interaction observability, which will influence the consistency of private reputations; and (iii) empathy when judging others, which again could influence the consistency of reputations.

Another potential use of reputation is **partner-choice**. As in IR models, individuals gain a good reputation by helping others, but in this framework, that reputation is then used to inform decisions about the selection of partners. These models are often labelled **Competitive Altruism** because competition to be selected by potential partners can lead to an escalation of helping behaviour (Roberts, 1998; Barclay, 2004, 2010, 2011, 2012, 2016; Hardy & Van Vugt, 2006; Barclay & Willer, 2007; Hardy & Van Vugt, 2006; Van Vugt et al., 2007; Sylwester & Roberts, 2010; Macfarlan et al., 2012; Böhm & Regner, 2013; Roberts, 2015; Herrmann et al., 2019). *Roberts et al*. offer a comparative perspective on models of indirect reciprocity and reputation-based partner choice, by analysing their assumptions. This comparison shows that despite superficial similarities, such as the success of both models in describing how reputation can promote cooperation, these models use vastly different assumptions. While IR models assume one-off interactions between randomly chosen individuals, reputation-based partner choice assumes active partner choice. The authors argue that this latter scenario is more relevant in interactions among humans, thus reputation-based partner choice models deserve more attention.

Reputation systems can be used in different contexts thus the **content of reputations** may greatly vary. While most models investigate a simplified system of "good" vs. "bad" reputation (see IR), real life reputation systems are more complex (Dukerich, Carter, 2000; Heugens, van Riel, van den Bosch, 2004; Fombrun, 1996; Rhee, Haunschild, 2006; Macfarlan et al., 2013; Dafoe, Renshon, Huth, 2014; Bartley & Child, 2014; Tadelis, 2016; Rogers, Goldstein, Fox 2018; Barclay & Barker, 2020). *Romano et al*. investigate how differences across social ecologies, institutions, and cultures influence the workings of reputation systems. Their interdisciplinary review also contrasts empirical insights from various disciplines about ecological factors that shape reputations with the way how models address the questions of reputation-based cooperation. *Garfield et al.* use the Human Relations Area Files ethnographic database to characterize the cross-cultural diversity in reputation domains. They find that while there is substantial variation (not only between societies but also between men and women), reputations for cultural conformity, prosociality, social status, and neural capital are widespread.

Another important question is whether the **everyday use of gossip** is consistent with the leading theories of reputation and cooperation. *Dores Cruz et al.* test several predictions from theories of indirect reciprocity and reputation-based partner selection about the content of everyday gossip. Using experience sampling methods, the researchers asked a Dutch community sample to report on recent gossip events they experienced in order to acquire information about the content and the consequences of gossip in everyday life. Their results lend support to the hypothesis that gossip is

often about other's cooperativeness and that it has consequences on the target's reputation and the receiver's intentions to interact with them in the future. *Hess and Hagen* used experimental and survey vignettes among MTurkers and California sorority women to investigate how the social context of gossip impacts its content and its perceived consequences. They found that people were more likely to gossip, and to share negative gossip, when situations were more competitive, and that having allies may help militate against negative gossip.

A crucial ingredient of aggregated reputation systems is **communication**. While communication is an important part of aggregated reputation systems, humans can communicate about a wide variety of things. **Signalling theory** investigates the conditions of deceptive vs. honest (informative) interactions between signallers and receivers (Spence, 1973; Grafen, 1990; Maynard Smith & Harper, 2003). Signals may carry information about intentions, conditions, or quality (Maynard Smith and Harper, 2003; Searcy and Nowicki, 2005; Bird et al., 2005; 2018; Barker et al., 2019). Seminal signalling models investigate **signals of quality** (Grafen, 1990, Godfray, 1991; Maynard Smith, 1991) - these models often named as "**honest signalling**" models- investigate the conditions when observable signals correlate with non-observable (but relevant) qualities (e.g., mate value). *Dumas et al.* investigate how additional information about an individual's social prominence or social capital can alter the signalling system, by using analytical and agent-based models. The inclusion of this information can generate a „reputational shield" and a „reputational poverty trap"; where a "reputational shield," involves low quality individuals being mistaken as high quality due to their social standing, and a "reputational poverty trap," implies that high quality individuals are unable to improve their reputation due to a lack of social capital. Drawing on a case study of religious signalling in South India, they suggest that more attention should be paid to the consequences of the complex feedbacks between reputation, prominence, and social capital.

Another important issue of signalling theory is the evolution of **conventional signals** (Dawkins & Guilford, 1995). The correlation between meaning and form is an emergent property of conventional signalling systems; the same form might have different meanings in different signalling systems. **Ritualization** is often proposed as a mechanism to explain the evolution of conventional signals. Unlike in other animals, in humans rituals can gain normative force through individuals' consensual understanding that these behaviours ought to be followed because they constitute markers of group distinction. *Przepiorka and Diekmann* call such norms signalling norms and investigate the conditions under which they can emerge. They argue it is the uncertainty about in-group members' trustworthiness that drives the emergence of signalling norms (Fehrler and Przepiorka 2013; Posner 2000; Przepiorka and Diekmann, 2013). As they identify the conditions for the emergence of seemingly wasteful and even harmful behaviours, they find that it is not the quality-revealing aspect but the in-group commitment such behaviours produce that stabilizes these very behaviours. That is, signalling norms promote parochial cooperation also because they impede cooperation with out-group members (Aksoy and Gambetta, 2016; Iannaccone, 1994; Sosis, 2005).

Crucially, humans are able to communicate about individuals who are not present. Evaluative communication of past experiences and observations about an absent third party is called **gossip**. A wealth of empirical studies shows that the "**threat of gossip**" can promote cooperative behaviour (Sommerfeld et al., 2007; Ellingsen, Johannesson, 2008; Piazza & Bering, 2008; Beersma & Gerben, 2011; Beersma & Van Kleef, 2012; Feinberg et al., 2012, 2014; Wu et al., 2015, 2016b; Samu et al., 2020). This threat of gossip refers to situations where individuals know that they are being observed and that these observations can be communicated to other group members. A general conclusion is that individuals will behave more cooperatively under the threat of gossip than they would do

otherwise. This clearly shows that humans are aware of the importance of reputations and of the importance of gossip as a mediator in reputation building in systems of aggregated reputation.

Gossip is often viewed as a cheap tool to maintain cooperation, but gossip alone will not maintain cooperation in the long term. The threat of gossip can promote a surge of cooperation, but such cooperation will collapse if the threat of gossip is not backed up by efficient sanctions against non-cooperators (Samu et al., 2020). This threat may then be ineffective if efficient sanctions are not possible, or if gossip cannot consistently inform individuals of who was cooperative and who was not. This raises the issue of the **reliability of gossip**. While most models assume that gossip is honest by default (i.e., individuals will communicate their experiences and observations with perfect fidelity), this need not be the case. There is rich evidence from different disciplines that humans are strategic and even manipulative in their communication (DePaulo et al., 1996; Kashy & DePaulo, 1996; Mann et al., 2014). We may not lie as much as we could (see Jacobsen et al., 2018 for review), but it does not change the fact that gossip may not always be honest (Buss & Dedden, 1990; Diekmann et al., 2014; Hess & Hagen, 2006; McAndrew, 2014; Fonseca & Peters, 2018). Studies show that widespread dishonest gossip could result in the collapse of reliable reputations, and thus the collapse of cooperation (Számadó et al., 2016).

So, what maintains the **honesty of gossip**? Several scholars took up the challenge to investigate the conditions which promote honest gossip. *Fonseca and Peters* investigated the ability of people to **detect** dishonest gossip and the willingness to reward those who spread honest gossip in an experimental Trust game. They found that people are reasonably adept at detecting dishonest gossip, yet the valence of gossip may matter for how gossipers are treated. People are less willing to trust gossipers who share negative gossip, even if this gossip is honest or ultimately helpful. In a complementary study, *Giardini et al*. investigated the effect of partner choice on gossip and cooperation in a sequential Public Goods Game. Introducing partner choice increased both the frequency of cooperative choices and of honest gossip among cooperators. When participants competed to be chosen there was a correlation between cooperation and honesty, thus suggesting that the opportunity for partner-choice promotes both the honesty of gossip and cooperation. The way in which gossip was reported was also different between co-operators and defectors who preferred to pass information as coming from someone else as a way of distancing from the source and thus avoid retaliation. *Samu and Takács* studied cooperative behaviour in a laboratory experiment where subjects played two-person Prisoner's Dilemma games with varying interaction partners and were allowed to gossip about their experiences. They investigated the role of two mechanisms that could support the honesty and credibility of gossip: they tested how cross-checking the validity of third-party information and allowing for social bonding between the sender and the receiver affected the reliability of reputations, and ultimately cooperation. They found that, while overall individuals trusted gossip, neither cross-checking or bonding opportunities increased cooperation. However, when there was a stronger competition for reputation, this tended to increase cooperation. *Wu et al.* use fitness interdependence theory (or stake, Roberts, 2005; Brown and Brown, 2006; Aktipis et al., 2018; Cronk et al., 2019) to model individuals' gossip strategies (honest vs. dishonest gossip) in a gossip triad across situations where the gossiper has varying levels of fitness interdependence with the target and/or the receiver. They investigate gossip strategies across four seminal games (i.e., stag-hunt game, snowdrift game, helping game, and punishment game) to illustrate how the gossiper's action is determined by (a) the gossiper's fitness interdependence with the receiver and the target, and (b) the marginal cost/benefit in terms of payoff differences between two possible game actions for the receiver and the target (i.e., game type). These models suggest a simple 'matching rule' that gossipers can use to make optimal decisions even under noise: be honest when there is a perfect match between the valence of fitness

interdependences and effects of honest gossip on receiver and target; be dishonest when there is a perfect mismatch between these two factors.

While most models of reputation (e.g., most IR models) assume a so called well-mixed population, where individuals can interact with any other member of the population, this is not a realistic assumption for human groups. While the role of **population structure** (i.e., spatially explicit models, network topology, network dynamics) is well investigated in two-person social dilemma games (e.g., Prisoner's Dilemma, Snowdrift game, Nowak, M. A., & May, 1992, 1993; Hauert, 2002; Ohtsuki et al., 2006; Számadó et al., 2008; see for reviews: Szabó & Fáth, 2007; Perc & Szolnoki, 2010) this issue has received less attention in the study of reputation systems. *Takács et al.* review how the structure and dynamics of interactions help or prevent the establishment and sustainability of reputation-based cooperation. They review both theoretical and empirical work on how fixed topologies and co-evolving networks can facilitate cooperation, how reputations are formed if interactions take place in a social network, and how these mechanisms impact cooperation in small- and large-scale societies. Besides the overview on how networks act as catalysts of cooperation, they also highlight how certain network topologies can lead to the dissemination of biased beliefs and false reputations, and limit cooperation to in-group encounters but foster conflicts with out-groups.

Punishment is an important mechanism that can promote cooperation (Yamagishi, 1986; Fehr & Gachter, 2000; 2002; Balliet et al., 2011; Ostrom, 2012). Punishment can be part of a reputation system, but punishment can also maintain cooperation without reputations (i.e., based on experience or direct observation). *Molho and Wu* investigate the differences between direct punishment tactics (i.e., physical and verbal confrontation) and indirect reputation-based tactics (i.e., gossip and ostracism). They describe the costs, in terms of potential retaliation, and benefits, in terms of recalibrating others' behaviour, of direct and indirect punishment tactics. Further, they review the different functions of punishment and reputation-based tactics, as well as their antecedents. Finally, they highlight the need to better understand how distinct tactics to intervene against offenses impact reputations. *Podder et al.* investigated cooperation problems in which individuals have a choice to opt out of the interaction (the optional Public Goods Game). The assignment of reputations is more complex in these interactions as there is more than a collectively good (cooperation) and bad (defection) choice. They find that the social norm that assigns a more moderate reputational penalty for opting out than for defection is able to sustain cooperation. Furthermore, they investigate the role of reputational dynamics in a setting where opting out as well as punishment can simultaneously be present and find that the two institutions work synergistically and promote large-scale cooperation under social norms that do not punish opting out choices to harshly.

For cooperation to persist, cooperators, on average, must be better-off than non-cooperators. *Tsvetkova* investigates how public and objective reputational information affects **payoff inequality** in repeated social dilemma interactions in large groups with the use of an agent-based model with simple decision heuristics and varying partner matching rules. She demonstrates that reputational information does not necessarily increase inequality in strategically updated networks, and actually decreases it in randomly rewired networks. More importantly, reputational information almost always results in high correlation between final payoffs and cooperative behaviour. Reputational information may therefore both improve cooperation and reduce inequality. These model results are validated with data from cooperation experiments showing similar correlations.

Another important ingredient of cooperation is **trust**. Trust promotes cooperation, but organisms often need to know how much to trust others to cooperate. Fitness interdependence (or stake, Roberts, 2005; Brown and Brown, 2006; Aktipis et al., 2018; Cronk et al., 2019) can promote

cooperation. *Barclay et al*. investigate whether individuals are able to find out the level of fitness interdependence from the level of helping behaviour. In their mathematical model, agents condition their help towards others based on their stake in the recipient's welfare, and recipients use other's helping decisions to assess whom to trust. They demonstrate that recipients will demand signals from others when they assign less value to the signallers and under conditions involving a higher cost of betrayed trust. Signal costs, however, are higher when signallers have a stronger incentive to defect.

4. Concluding remarks

While reputation has been claimed to be one of the key drivers of the success of human large-scale cooperation, how humans achieve this success is less well understood. Humans are uniquely able to enjoy the benefits of reputation-based cooperation, due to our advanced abilities for communication, and expansive cognitive and social skills (e.g., working memory and perspective taking). The theme issue outlines promising research directions about the evolution and conditions of reputation systems, the context and content of reputation, and the operation of reputation-based systems in social networks.

With this theme issue and the workshop that originally sparked it, we have sought to bring together the different strands of research on reputation-based cooperation being carried out across the diversity of the behavioral sciences. The authors of the pieces in this issue are housed in different departments so inevitably bring to this field distinct sets of questions, theories, and tools of inquiry. Indeed, if one were to survey the operationalizations of core concepts such as reputation, gossip, punishment, or cooperation in the papers in this theme issue, slight differences could surely be found. Despite this, we have found there to be important areas of shared understanding and interest, across these traditional divides. There is a shared foundation and a shared recognition that more work is needed to advance the field: to add necessary complexity, whether to our models or our experiments, and to expand our empirical evidence, whether through investigations in new and more diverse settings, or through new observational tools and data sources.

The theme issue also highlights that our knowledge is still fragmented and there is no unified framework that could outline the necessary conditions for honest gossip that could create and maintain reliable reputations. While models about reputation systems are useful, they do not always reflect human social networks and reputation systems. For example, empirical investigations show that reputation is often more complex than the assignment of a binary evaluation. Another example is that communication need to be to a large extent honest to sustain reputation-based systems in these models yet empirical studies show that gossip need not be always honest in real life situations. Network structure can also influence the flow of communication and thus the level of cooperation.

Further interdisciplinary work is much needed to resolve the issues identified above. The success of human cooperation is strongly supported by reputation systems, yet we are just about to understand the complexity of these systems. While this theme issue offers many novel insights it identifies just as many open issues. We have still much to learn about the language of cooperation.

Acknowledgements

References

Aktipis, A., Cronk, L., Alcock, J., Ayers, J. D., Baciu, C., Balliet, D., Boddy, A. M., Curry, O. S., Krems, J. A., Muñoz, A., Sullivan, D., Sznycer, D., Wilkinson, G. S., & Winfrey, P. (2018). Understanding cooperation through fitness interdependence. Nature Human Behaviour, 2(7), 429–431. https://doi.org/10.1038/s41562-018-0378-4

Aksoy, O., & Gambetta, D. (2016). Behind the Veil: The Strategic Use of Religious Garb. European Sociological Review, 6(1), 792-806. doi:10.1093/esr/jcw035

Balliet, D., Mulder, L. B., & Van Lange, P. A. (2011). Reward, punishment, and cooperation: a meta-analysis. Psychological Bulletin, 137(4), 594.

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

Barclay, P. (2010). Altruism as a courtship display: Some effects of third-party generosity on audience perceptions. British Journal of Psychology, 101(1), 123-135.

Barclay, P. (2011). Competitive helping increases with the size of biological markets and invades defection. Journal of Theoretical Biology, 281(1), 47-55.

Barclay, P. (2012). Harnessing the Power of Reputation: Strengths and Limits for Promoting Cooperative Behaviors. Evolutionary Psychology, 10(5), 147470491201000500.

Barclay, P. (2016). Biological markets and the effects of partner choice on cooperation and friendship. Current Opinion in Psychology, 7, 33-38.

Barclay, P., & Barker, J. L. (2020). Greener Than Thou: People who protect the environment are more cooperative, compete to be environmental, and benefit from reputation. Journal of Environmental Psychology, 72, 101441.

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

Barclay, P., Bliege Bird, R., Roberts, G., Számadó, Sz. (In press) Cooperating to show that you care: costly helping as an honest signal of fitness interdependence. Philos. Trans. R. Soc. B Biol. Sci.

Barrett, H.; Saxe, R. (In press) Explaining variation in the role of intentions in moral judgment. Philos. Trans. R. Soc. B Biol. Sci.

Barker, J. L., Power, E. A., Heap, S., Puurtinen, M., & Sosis, R. (2019). Content, cost, and context: A framework for understanding human signaling systems. Evolutionary Anthropology: Issues, News, and Reviews, 28(2), 86–99. https://doi.org/10.1002/evan.21768

Bartley, T., & Child, C. (2014). Shaming the corporation: The social production of targets and the anti-sweatshop movement. American Sociological Review, 79(4), 653-679.

Beersma B, Gerben AVK (2011) How the Grapevine Keeps You in Line: Gossip Increases Contributions to the Group. Soc Psychol and Pers 2011, DOI: 10.1177/1948550611405073.

Beersma, B., & Van Kleef, G. A. (2012). Why people gossip: An empirical analysis of social motives, antecedents, and consequences. Journal of Applied Social Psychology, 42(11), 2640-2670.

Bird, R.B., & Smith, Eric A. (2005). Signaling Theory, Strategic Interaction, and Symbolic Capital. Current Anthropology, 46(2), 221-248.

Bird, R. B., Ready, E., & Power, E. A. (2018). The social significance of subtle signals. Nature Human Behaviour, 2(7), 452-457.

Böhm, R., & Regner, T. (2013). Charitable giving among females and males: an empirical test of the competitive altruism hypothesis. Journal of Bioeconomics, 15(3), 251-267.

Brown SL, Michael Brown R. 2006 Selective investment theory: Recasting the functional significance of close relationships. Psychol. Inq. 17, 1–29. (doi:10.1207/s15327965pli1701_01)

Bshary, R., & Grutter, A. S. (2006). Image scoring and cooperation in a cleaner fish mutualism. Nature, 441(7096), 975-978.

Buss, D. M., & Dedden, L. A. (1990). Derogation of competitors. Journal of Social and Personal Relationships, 7(3), 395–422.

Cronk L, Steklis D, Steklis N, van den Akker OR, Aktipis A. 2019 Kin terms and fitness interdependence. Evol. Hum. Behav. 40, 281–291. (doi:10.1016/j.evolhumbehav.2018.12.004)

Dafoe, A., Renshon, J., & Huth, P. (2014). Reputation and status as motives for war. Annual Review of Political Science, 17, 371-393.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. Journal of personality and social psychology, 70(5), 979.

Diekmann, A., Jann, B., Przepiorka, W., & Wehrli, S. (2014). Reputation Formation and the Evolution of Cooperation in Anonymous Online Markets. American Sociological Review, 79(1), 65–85.

Dores Cruz, T., Thielmann, I., Columbus, S., Molho, C., Wu, J., Righetti, F., De Vries, R., Koutsoumpis, A., Van Lange, P., Beersma, B., Balliet, D., (In press) Gossip and Reputation in Everyday Life. Philos. Trans. R. Soc. B Biol. Sci.

Dukerich, J. M., & Carter, S. M. (2000). Distorted images and reputation repair. The expressive organization: Linking identity, reputation, and the corporate brand, 97-112.

Dumas, M., Power, E., Barker, J., (In press) When does reputation lie? Dynamic feedbacks between costly signals, social capital, and social prominence. Philos. Trans. R. Soc. B Biol. Sci.

Ellingsen T, Johannesson M. (2008). Anticipated verbal feedback induces altruistic behavior. Evol and Human Behav 29, 100–105.

Fehrler, S., & Przepiorka, W. (2013). Charitable giving as a signal of trustworthiness: Disentangling the signaling benefits of altruistic acts. Evolution and Human Behavior, 34(2), 139-145. doi:10.1016/j.evolhumbehav.2012.11.005

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. American Economic Review, 90(4), 980-994.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137-140.

Feinberg, M., Willer, R., Stellar, J., & Keltner, D. (2012). The virtues of gossip: reputational information sharing as prosocial behavior. Journal of personality and social psychology, 102(5), 1015.

Feinberg, M., Willer, R., & Schultz, M. (2014). Gossip and ostracism promote cooperation in groups. Psychological science, 25(3), 656-664.

Fombrun, C. J., & Rindova, V. (1996). Who's tops and who decides? The social construction of corporate reputations. New York University, Stern School of Business, Working Paper, 5-13.

Fonseca MA, Peters K. 2018 Will any gossip do? Gossip does not need to be perfectly accurate to promote trust. Games Econ. Behav. 107, 253–281. (doi:10.1016/j.geb.2017.09.015)

Fonseca, M., Peters, K., (In press) Is It Costly to Deceive? People Are Adept at Detecting Gossipers' Lies but May Not Reward Honesty. Philos. Trans. R. Soc. B Biol. Sci.

Garfield, Z., Schacht, R., Post, E., Uehling, A., Ingram, D., MacFarlan, S., (In press) The content and structure of reputation domains across human societies: a view from the evolutionary social sciences. Philos. Trans. R. Soc. B Biol. Sci.

Giardini, F., & Wittek, R. (2019b). Gossip, Reputation, and Sustainable Cooperation. Sociological Foundations. In Giardini, F., & Wittek, R. (Eds.). (2019). The Oxford Handbook of Gossip and Reputation. Oxford Handbooks, 23-46.

Giardini, F., Vilone, D., Sánchez, A., Antonioni, A., (In press) Gossip and competitive altruism support cooperation in a Public Good Game. Philos. Trans. R. Soc. B Biol. Sci.

Godfray HC. 1991 Signalling of need by offspring to their parents. Nature 352. (doi:10.1038/352328a0)

Grafen A. 1990 Biological signals as handicaps. J. Theor. Biol. 144, 517–546. (doi:10.1016/S0022-5193(05)80088-8)

Granovetter M (1978) Threshold Models of Collective Behavior. American Journal of Sociology, 83, 489–515.

Guilford, T., & Dawkins, M. S. (1995). What are conventional signals?. Animal Behaviour, 49(6), 1689-1695.

Hardy, C. L., & Van Vugt, M. (2006). Nice guys finish first: The competitive altruism hypothesis. Personality and Social Psychology Bulletin, 32(10), 1402-1413.

Hauert, C. (2002). Effects of space in 2× 2 games. International Journal of Bifurcation and Chaos, 12(07), 1531-1548.

Hess, N. H., & Hagen, E. H. (2006). Sex differences in indirect aggression: Psychological evidence from young adults. Evolution and Human Behavior, 27(3), 231–245.

Hess, N., Hagen, E., (In press) Competitive gossip: The impact of domain, resource value, resource scarcity, and coalitions. Philos. Trans. R. Soc. B Biol. Sci.

Herrmann, E., Engelmann, J. M., & Tomasello, M. (2019). Children engage in competitive altruism. Journal of Experimental Child Psychology, 179, 176-189.

Heugens, P. P., Van Riel, C. B., & Van Den Bosch, F. A. (2004). Reputation management capabilities as decision rules. Journal of Management Studies, 41(8), 1349-1377.

Iannaccone, L. R. (1994). Why Strict Churches Are Strong. American Journal of Sociology, 99(5), 1180-1211.

Kashy, D. A., & DePaulo, B. M. (1996). Who lies?. Journal of Personality and Social Psychology, 70(5), 1037.

Kollock, P. (1998). Social dilemmas: The anatomy of cooperation. Annual review of sociology, 24(1), 183-214.

Leimar, O., & Hammerstein, P. (2001). Evolution of cooperation through indirect reciprocity. Proceedings of the Royal Society of London. Series B: Biological Sciences, 268(1468), 745–753.

Macfarlan, S. J., Remiker, M., & Quinlan, R. (2012). Competitive Altruism Explains Labor Exchange Variation in a Dominican Community. Current Anthropology, 53(1), 118-124.

Macfarlan, S. J., Quinlan, R., & Remiker, M. (2013). Cooperative behaviour and prosocial reputation dynamics in a Dominican village. Proceedings of the Royal Society of London B: Biological Sciences, 280(1761), 20130557.

Mann, H., Garcia-Rada, X., Houser, D., & Ariely, D. (2014). Everybody else is doing it: Exploring social transmission of lying behavior. PloS one, 9(10), e109591.

Manrique, H.M., Zeidler, H., Roberts, G., Barclay, P., Walker, M., Samu, F., Farina, A., Bshary, R., Raihani, N., (In press) The psychological foundations of reputation-based cooperation. Philos. Trans. R. Soc. B Biol. Sci.

Marsh, C. (2018). Indirect reciprocity and reputation management: Interdisciplinary findings from evolutionary biology and economics. Public Relations Review, 44(4), 463-470.

Maynard Smith, J (1991) Honest signalling: the Philip Sidney game. Animal Behaviour, 42(6):1034-1035.

Maynard Smith J, Szathmáry E (1995) The Major Transitions in Evolution. Oxford.

Maynard Smith, J., Harper, D. (2003). Animal signals. Oxford University Press.

McAndrew, F. T. (2014). The "sword of a woman": Gossip and female aggression. Aggression and Violent Behavior, 19(3), 196–199.

Milinski, M., Semmann, D., & Krambeck, H. J. (2002). Reputation helps solve the 'tragedy of the commons'. Nature, 415(6870), 424.

Milinski, M. (2016). Reputation, a universal currency for human social interactions. Phil. Trans. R.Soc. B, 371(1687), 20150100.

Molho, C., Wu, J., (In press) Direct Punishment and Indirect Reputation-Based Tactics to Intervene Against Offenses. Philos. Trans. R. Soc. B Biol. Sci.

Nowak, M. A., & May, R. M. (1992). Evolutionary games and spatial chaos. Nature, 359(6398), 826-829.

Nowak, M. A., & May, R. M. (1993). The spatial dilemmas of evolution. International Journal of bifurcation and chaos, 3(01), 35-78.

Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. Nature, 393(6685), 573.

Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. Nature, 437(7063), 1291.

Noë, R., & Hammerstein, P. (1995). Biological markets. Trends in Ecology & Evolution, 10(8), 336-339.

Ohtsuki, H., & Iwasa, Y. (2004). How should we define goodness? Reputation dynamics in indirect reciprocity. Journal of Theoretical Biology, 231(1), 107-120.Ohtsuki, H., & Iwasa, Y. (2006). The leading eight: Social norms that can maintain cooperation by indirect reciprocity. Journal of Theoretical Biology, 239(4), 435-444.

Ohtsuki, H., Hauert, C., Lieberman, E., & Nowak, M. A. (2006). A simple rule for the evolution of cooperation on graphs and social networks. Nature, 441(7092), 502-505.

Okada, I. (2020). A review of theoretical studies on indirect reciprocity. Games, 11(3), 27.

Ostrom, E. (2012). Experiments combining communication with punishment options demonstrate how individuals can overcome social dilemmas. Behavioral and Brain Sciences, 35(1), 33.

Pacheco, J. M., Santos, F. C., & Chalub, F. (2006). Stern-judging: A simple, successful norm which promotes cooperation under indirect reciprocity. Plos Computational Biology, 2(12), 1634-1638.

Panchanathan, K., & Boyd, R. (2004). Indirect reciprocity can stabilize cooperation without the second-order free rider problem. Nature, 432(7016), 499.

Perc, M., & Szolnoki, A. (2010). Coevolutionary games—a mini review. BioSystems, 99(2), 109-125.

Peters, K., & Fonseca, M. A. (2020). Truth, lies, and gossip. Psychological Science, 31(6), 702-714.

Piazza J, Bering JM. (2008). Concerns about reputation via gossip promote generous allocations in an economic game. Evol and Human Behav 2008, 29, 172–178.

Podder, S., Righi, S. & Takács, K. (2021). Local reputation, local selection, and the leading eight norms. *Sci Rep* **11,** 16560.

Podder, S., Righi, S., Pancotto, F., (In press) Reputation and Punishment sustain cooperation in the Optional Public Goods Game. Philos. Trans. R. Soc. B Biol. Sci.

Posner, E. A. (2000). Law and Social Norms. Cambridge (MA): Harvard University Press.

Przepiorka, W., & Diekmann, A. (2013). Temporal Embeddedness and Signals of Trustworthiness: Experimental Tests of a Game Theoretic Model in the United Kingdom, Russia, and Switzerland. European Sociological Review, 29(5), 1010-1023. doi:10.1093/esr/jcs078

Przepiorka, W., Diekmann, A., (In press) Parochial cooperation and the emergence of signalling norms. Philos. Trans. R. Soc. B Biol. Sci.

Rhee, M., & Haunschild, P. R. (2006). The liability of good reputation: A study of product recalls in the US automobile industry. Organization Science, 17(1), 101-117.

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.

Roberts, G. (2005). Cooperation through interdependence. Animal Behaviour, 70, 901-908.

Roberts, G. (2008). Evolution of direct and indirect reciprocity. Proceedings of the Royal Society B-Biological Sciences, 275(1631), 173-179.

Roberts, G. (2015). Human Cooperation: The Race to Give. Current Biology, 25(10), R425-R427.

Roberts, G., Raihani, N., Bshary, R., Manrique, H.M., Farina, A., Samu, F., Barclay, P., (In press) The benefits of being seen to help others: indirect reciprocity and reputation-based partner choice. Philos. Trans. R. Soc. B Biol. Sci.

Rogers, T., Goldstein, N. J., & Fox, C. R. (2018). Social Mobilization. Annual Review of Psychology, 69, 357-381.

Romano, A., Giardini, F., Columbus, S., De Kwaadsteniet, E., Kisfalusi, D., Triki, Z., Snijders, C., Hagel, K., (In press) Reputation and Socio-Ecology in Humans. Philos. Trans. R. Soc. B Biol. Sci.

Samu F, Számadó S, Takács K. (2020). Scarce and directly beneficial reputations support cooperation. Sci. Rep. 10, 11486. (doi:10.1038/s41598-020-68123-x)

Samu, F., Takács, K., (In press) Evaluating mechanisms that could support credible reputations and cooperation: cross-checking and social bonding. Philos. Trans. R. Soc. B Biol. Sci.

Santos, F. P., Santos, F. C., & Pacheco, J. M. (2018). Social norm complexity and past reputations in the evolution of cooperation. Nature, 555, 242.

Santos, F.P., Pacheco, J., Santos, F., (In press) The complexity of human cooperation under indirect reciprocity. Philos. Trans. R. Soc. B Biol. Sci.

Searcy, W. A. & Nowicki, S. (2005). The Evolution of Animal Communication. Reliability and Deception in Signaling Systems. Princeton University Press, Princeton.

Shank, D. B., Kashima, Y., Peters, K., Li, Y., Robins, G., & Kirley, M. (2019). Norm talk and human cooperation: Can we talk ourselves into cooperation?. Journal of Personality and Social Psychology, 117(1), 99.

Sigmund, K. (2012). Moral assessment in indirect reciprocity. Journal of Theoretical Biology, 299, 25-30.

Simpson, B., & Willer, R. (2015). Beyond altruism: Sociological foundations of cooperation and prosocial behavior. Annual Review of Sociology, 41, 43-63.

Smith, E. A. (2010). Communication and collective action: language and the evolution of human cooperation. Evolution and Human Behavior, 31(4), 231-245.

Sosis, R. (2005). Does Religion Promote Trust? The Role of Signaling, Reputation, and Punishment. Interdisciplinary Journal of Research on Religion, 1, 1-30.

Sommerfeld, R. D., Krambeck, H. J., Semmann, D., & Milinski, M. (2007). Gossip as an alternative for direct observation in games of indirect reciprocity. Proceedings of the National Academy of Sciences, 104(44), 17435-17440.

Spence, M. (1973). Job market signaling. The Quarterly Journal of Economics 87, 355–374.

Szabó, G., & Fath, G. (2007). Evolutionary games on graphs. Physics Reports, 446(4-6), 97-216.

Számadó, S., Szalai, F., & Scheuring, I. (2008). The effect of dispersal and neighbourhood in games of cooperation. Journal of Theoretical Biology, 253(2), 221-227.

Számadó, S., Szalai, F., & Scheuring, I. (2016). Deception undermines the stability of cooperation in games of indirect reciprocity. PloS One, 11(1), e0147623.

Sylwester, K., & Roberts, G. (2010). Cooperators benefit through reputation-based partner choice in economic games. Biology Letters, 6(5), 659-662.

Takács, K., Gross, J., Testori, M., Letina, S., Kenny, A., Power, E., Wittek, R., (In press) Networks of Reliable Reputations and Cooperation: A Review. Philos. Trans. R. Soc. B Biol. Sci.

Tadelis, S. (2016). Reputation and feedback systems in online platform markets. Annual Review of Economics, 8, 321-340.

Tennie, C., Frith, U., & Frith, C. D. (2010). Reputation management in the age of the world-wide web. Trends in cognitive sciences, 14(11), 482-488.

Tsvetkova, M., (In press) The effects of reputation on inequality in network cooperation games. Philos. Trans. R. Soc. B Biol. Sci.

Van Vugt, M., Roberts, G., & Hardy, C. (2007). Competitive altruism: Development of reputation-based cooperation in groups. In R. Dunbar & L. Barrett (Eds.), Handbook of Evolutionary Psychology. Oxford University Press.

Wu, J., Balliet, D., & Van Lange, P. A. (2015). When does gossip promote generosity? Indirect reciprocity under the shadow of the future. Social Psychological and Personality Science, 6(8), 923-930.

Wu, J., Balliet, D., & Van Lange, P. A. (2016a). Gossip versus punishment: The efficiency of reputation to promote and maintain cooperation. Scientific Reports, 6, 23919.

Wu, J., Balliet, D., & Van Lange, P. A. (2016b). Reputation management: Why and how gossip enhances generosity. Evolution and Human Behavior, 37(3), 193-201.

Wu, J., Számadó, Sz., Barclay, P., Beersma, B., Dores C., Terence; Lo Iacono, S., Nieper, A.S., Peters, K., Przepiorka, W., Tiokhin, L., Van Lange, P., (In press) Honesty and Dishonesty in Gossip Strategies: A Fitness Interdependence Analysis. Philos. Trans. R. Soc. B Biol. Sci.

Yamagishi, T. (1986). The provision of a sanctioning system as a public good. Journal of Personality and social Psychology, 51(1), 110.