

## 6 Why We Cooperate

Jillian Jordan, Alexander Peysakhovich, and David G. Rand

What does it mean to be moral? In general, answers to this question can take a normative (“what should be”) or positive (“what is”) perspective. Traditionally, philosophers have focused on normative answers to this question, whereas psychologists, evolutionary biologists, and social scientists have focused on the positive aspects of morality: How and why did our sense of morality evolve? What psychological processes contribute to our moral judgments? Which moral rules are universal, and which vary across cultures? In this chapter, we ask these positive questions about a particular set of moral behaviors: *cooperative* behaviors.

We define cooperative (or prosocial) behaviors as cases in which an individual pays a personal cost to provide a benefit to another individual or group of individuals. We broadly define *cost*, considering a wide range of potential costs including time and effort, resources and money, and physical harm. In most of the cases that we consider, the cost of cooperation is smaller than the benefit it creates, making cooperation productive: pairs or groups of individuals who can successfully cooperate with each other are better off than those who cannot. Because of the benefits cooperation creates, many people consider there to be a moral obligation to cooperate.

Cooperation is ubiquitous in the world around us, at a large scale with humanitarian organizations such as the Red Cross to conservation groups such as the World Wildlife Fund and on a small scale with people choosing to cover shifts for sick co-workers or to help friends move from one apartment to another. Although a gene’s-eye view of evolution can explain cooperation among relatives, cooperation between unrelated individuals poses a puzzle from both the perspective of natural selection and that of rational self-interest. Why should individuals make sacrifices to help potential competitors succeed? What motivates such prosocial behavior?

In recent years many researchers have sought to shed light on these questions by using economic games to investigate cooperative behavior (for reviews, see Camerer, 2003; Rand & Nowak, 2013). In these studies game theoretical approaches are used to create experimental paradigms: people make choices that affect the money earned by themselves and other players. One such paradigm that has become particularly popular is the public goods game (PGG). In a typical PGG four people are placed into a group (typically interacting anonymously through a computer interface), and each is endowed with \$10. Each then chooses how much money to contribute to a “common project,” with all contributions being doubled and split evenly among the four group members. Under these rules if everyone contributes all of her or his endowment, everyone’s money doubles: cooperation is productive. However, individuals personally lose money on contributing, because for each \$1 contributed, the contributor receives back only \$0.50. Thus, one’s earnings are maximized by contributing nothing and “free riding” off the contributions of others. In this way the PGG captures the tensions between individual and collective interests that are the heart of the cooperation. Closely related to the PGG is the prisoner’s dilemma (PD), probably the most widely studied economic game. The PD is a two-person version of the PGG in which players make a binary choice between a socially optimal action (often defined as paying a cost to benefit the other player) or an individually optimal action. The PD, PGG, and other related games have been used in thousands of theoretical and experimental studies to understand when and why people cooperate. In this chapter we review key insights that have emerged from this work.

### **Mechanisms That Promote Cooperation**

In social dilemmas such as the PGG and PD, selfishness always earns more than cooperation (hence the dilemma). However, the one-shot anonymous interactions described by these games omit key elements of our world that can allow cooperators to outcompete noncooperators. We refer to alternative interaction structures that improve the long-run payoff of cooperating as mechanisms that promote cooperation (Nowak, 2006; Rand & Nowak, 2013). These mechanisms provide *ultimate* explanations for why people cooperate: if cooperation actually pays off, evolution and rational self-interest can both favor cooperating. We focus on mechanisms that can explain

cooperation among unrelated individuals rather than on theories related to kin selection because cooperation among nonkin is the major challenge facing modern human societies.

### **Direct Reciprocity**

Perhaps the most important mechanism for most human interactions is *direct reciprocity*: individuals interact repeatedly and condition their cooperative behavior on the cooperation of their partners (Axelrod, 1984; Fudenberg & Maskin, 1986). Under direct reciprocity the “shadow of the future” motivates individuals to cooperate today in order to receive the benefits of cooperation tomorrow. Direct reciprocity provides the basis for many long-term relationships in humans, such as friendships and work partnerships.

A simple strategy, “tit-for-tat” (TFT), captures the essence of direct reciprocity (Axelrod, 1984). An individual using the TFT strategy begins a repeated interaction by cooperating and then in subsequent interactions copies her partner’s last action. Thus, when playing against a TFT player, you have an incentive to cooperate now so that she will cooperate with you next time (provided the returns to cooperation are large enough relative to the chance that you wind up not meeting your partner again).

In most real cooperative interactions mistakes are possible: for example, forces beyond your control may prevent you from cooperating even when you want to, or you may be confused about which action is actually helpful to your partner. If both partners are using TFT, a single such mistake can derail an otherwise cooperative relationship. As a result the presence of errors favors reciprocal strategies that sometimes cooperate even after their partners have defected (Nowak & Sigmund, 1992; Rand, Ohtsuki, & Nowak, 2009).

Theoretical work on the evolution of cooperation in repeated games is complemented by experimental evidence that repetition can indeed lead to the emergence of cooperation. When subjects playing repeated PDs have a sufficiently large probability of interacting again with the same partner in the next round, they learn to cooperate (Blonski, Ockenfels, & Spagnolo, 2011; Dal Bó & Fréchette, 2011; Dreber, Rand, Fudenberg, & Nowak, 2008; Fudenberg, Rand, & Dreber, 2012; Rand, Fudenberg, & Dreber, 2013). Otherwise, when relationships are too short or the payoff for defection is too tempting, they learn to defect. Furthermore, introducing the possibility of mistakes causes players to adopt strategies that are lenient (i.e., that wait for multiple

defections before switching to defection themselves) and forgiving (i.e., that are willing to return to cooperation after periods of defection) (Dreber, Fudenberg, & Rand, 2014; Fudenberg et al., 2012; Rand et al., 2013).

### **Indirect Reciprocity**

Direct reciprocity can make cooperation between pairs of individuals advantageous. But what about cooperation at a larger scale than dyadic interactions? This question is answered in part by *indirect reciprocity*, whereby my actions toward you depend on your previous actions toward others (Nowak & Sigmund, 2005). Under indirect reciprocity people earn good reputations when they cooperate with others and thus can expect increased cooperation from future partners. Social norms within a community specify standards for acceptable behavior, and information about individuals' behavior is spread through gossip. Successful social norms often assign good reputations to those who cooperate with others in good reputation and defect with those in bad reputation. Thus, individuals with good reputations are then rewarded because they are more likely to be the recipients of cooperative behavior.

There are two distinct reasons why one might preferentially cooperate with individuals who are known to be cooperative. One might reason that individuals who are known to be cooperative are more likely to reciprocate cooperative behavior. Thus, individuals may discriminate based on reputation as a way to select desirable interaction partners (Barclay & Willer, 2007).

Alternatively, one might cooperate with other cooperators merely to maintain a good reputation and thus receive more cooperation in the future (Pfeiffer, Tran, Krumme, & Rand, 2012). Rather than using a partner's previous behavior as a signal about her future behavior, the partner's previous behavior stipulates what you must do in order to maintain a good reputation yourself. Although perhaps less intuitive, this latter logic is the overwhelming focus of theoretical work on indirect reciprocity.

Experimental work confirms the theorized importance of reputation in promoting human cooperation: people playing economic games learn to cooperate when it is sufficiently likely that others will know about their previous actions (Feinberg, Willer, Stellar, & Keltner, 2012; Milinski, Semmann, Bakker, & Krambeck, 2001; Pfeiffer et al., 2012; Wedekind & Milinski, 2000). Furthermore, evidence suggests that humans are so highly attuned to their reputations that even subtle images of eyespots can increase cooperation

by unconsciously priming the sense of being watched (Haley & Fessler, 2005; Sparks & Barclay, 2013). Reputation systems have also been shown to promote cooperation outside the laboratory: blood donation (Lacetera & Macis, 2010) and giving to charity (Karlan & McConnell, 2012) increase when donors' names are published, and people are three times as likely to sign up for an energy blackout reduction program when sign-ups are observable (Yoeli, Hoffman, Rand, & Nowak, 2013).

### **Institutions**

Scaling up even further, *institutions* provide an important tool for maintaining cooperation in large groups (Bowles, Choi, & Hopfensitz, 2003; North, 1990; Ostrom, 1990). Humans often explicitly design institutions for the purpose of incentivizing good behavior. For example, governments create criminal justice systems, often employing police and courts, to prevent anti-social behaviors such as theft and assault. Such legal institutions have a long history in human societies. Smaller organizations also employ formal codes of conduct and often designate specific individuals to enforce the rules.

Institutions may also refer to social structures that create infrastructure for cooperative exchanges, like markets. Markets provide a regulated environment for strangers to engage in productive trades of goods and services (Greif, 1993; Milgrom & North, 1990). In general, institutions can promote cooperation both by deterring bad behavior and by promoting trust that others will cooperate.

The role of institutions in human cooperation has received much less attention among experimentalists using economic games than direct and indirect reciprocity. But several recent studies have begun to demonstrate the power that institutional incentives have over human cooperative behavior in the lab (Andreoni & Gee, 2012; Baldassari & Grossman, 2011; Ouss & Peysakhovich, 2013). More work in this vein is an important direction for future research.

### **Proximate Psychology of Cooperation**

Because of mechanisms such as these, cooperation often pays off in the long run. Yet the suggestion that cooperation is only about maximizing long-term payoffs does not match well with our daily life experiences. A brief moment of introspection clearly indicates that all cooperative behavior

does not result from conscious calculations of expected returns. We are surrounded by examples of people who seem to help because they genuinely care for others. And almost all of us have acted altruistically at one time or another without considering future returns, only motivated by our notions of morality and ethical behavior. How can we explain this form of truly costly cooperation?

We argue that much of the explanation comes from distinguishing between ultimate and proximate explanations for cooperative behavior. *Ultimate* explanations describe *why* people cooperate by outlining how cooperation pays off in the long run. In contrast, *proximate* explanations can explain *how* people cooperate by outlining the motivations, emotions, and cognitions that lead to cooperation in the moment. Thus, when we say that an individual cooperates because of direct reciprocity, we are providing an ultimate explanation. In contrast, when we say that an individual cooperates because she genuinely cares for her friend, we are providing a proximate explanation. Ultimate explanations provide reasons why proximate mechanisms would have come to function as they do: it is long-run advantageous to cooperate with our friends, and so we come to care for our friends (as a psychological mechanism to help motivate us to cooperate). These types of proximate psychological mechanisms help us to make choices that are typically advantageous without having to engage in the costly effort of estimating future returns.

One of the most important proximate mechanisms for motivating cooperative behavior is empathy. The *empathy-altruism* hypothesis contends that *empathetic concern* is a human emotional response to taking the perspective of another person in need (Batson, Ahmad, & Lishner, 2011). Empathetic concern motivates humans to see that need relieved, often through cooperative helping. Across a large body of studies Batson and colleagues have showed that experimentally manipulating empathic concern increases cooperative behaviors, such as volunteering to take electric shocks instead of the target (Batson, Duncan, Ackerman, Buckley, & Birch, 1981) and cooperating with the target in economic games (Batson & Ahmad, 2001). Furthermore, evidence suggests that empathic individuals experience positive mood changes when they see a need get relieved, even if another agent caused the relief—suggesting that empathic concern reflects a genuine care for others (Batson et al., 1988).

### The Social Heuristics Hypothesis

Any complete theory of cooperation must explain why we often observe cooperative behavior when no ultimate mechanisms appear to be operative, or they appear too weak to incentivize cooperation. For example, in January 2010, over 3 million people text-messaged the American Red Cross to donate \$10 to disaster relief for the Haiti earthquake. It is hard to generate an ultimate explanation for such a behavior. Why would it pay off for individuals to give anonymous gifts to strangers in foreign countries? Likewise, in laboratory experiments, subjects routinely pay costs to benefit anonymous strangers in one-shot interactions in which they have nothing to gain (Camerer, 2003). Such “irrational” cooperation is quite common and is critical to the success of our societies and to our identities as moral actors.

The social heuristics hypothesis (SHH; Rand, Greene, & Nowak, 2012; Rand et al., 2014) provides one explanation for this intrinsically motivated cooperation. The SHH considers cooperative decisions from a *dual process* perspective, where decisions result from an interplay between two types of psychological processes: those that are fast, automatic, intuitive, and affective; and those that are slow, controlled, and deliberative (Kahneman, 2003; Sloman, 1996). The SHH contends that over time, strategies that are typically successful become automatized as default responses. As a result if cooperation typically pays off in the long run, individuals come to develop prosocial preferences and genuinely care for others. Individuals then carry these heuristics with them to other settings, such as one-shot anonymous laboratory experiments. Therefore, individuals who generally interact in environments where cooperation is advantageous should, as a result of spillover effects, be predisposed toward prosociality even when it will not actually pay off.

To provide evidence that cooperative behavior is the default response for many people, Rand et al. (2012) manipulated cognitive processing of subjects playing economic games. Across a series of studies PGG participants were randomly assigned to conditions that induced more intuitive decision making (via time pressure or writing about a time their intuitions had worked out well) or more deliberative decision making (via time delay or writing about a time that careful reasoning worked out well). The results showed significantly more cooperation among individuals induced to be intuitive than individuals induced to be deliberative. Thus, these results provided evidence that, on average, intuition favors cooperation in

one-shot anonymous social dilemmas. It has also been found that impairing the function of the right lateral prefrontal cortex, a brain region associated with deliberation and control, increases giving in a unilateral money transfer, whereas amplifying this region decreases giving (Ruff, Ugazio, & Fehr, 2013), and that people by default project a cooperative frame onto neutrally framed prisoner's dilemma games (Engel & Rand, 2014).

Does this mean that cooperation is automatic for all people at all times? The answer is no. Social heuristics are flexible, and they change as we learn through experience whether cooperation tends to pay off. Although Rand et al. (2012) found that intuition favored cooperation overall, this pattern applied only to subjects who reported having cooperative daily-life interaction partners. In contrast, subjects who did *not* report having cooperative partners showed no evidence of defaulting toward cooperation. Further evidence of the moderating role of trust on intuitive cooperation comes from Rand and Kraft-Todd (In press). Whereas intuitive responses varied based on life experience, deliberation favored selfishness among both trusting and nontrusting subjects. Additional evidence moderators come from Tinghög et al. (2013), who find no effect of time pressure on PGG cooperation across several studies; and from Rand et al. (2014), who find a variety of time-pressure effect sizes across fifteen different studies that range from large and positive to null but never significantly negative. Similar results, where intuition sometimes favors prosociality and sometimes has no effect, have been found using cognitive load (Cornelissen, Dewitte, & Warlop, 2011; Hauge, Brekke, Johansson, Johansson-Stenman, & Svedsäter, 2009; Roch, Lane, Samuelson, Allison, & Dent, 2000; Schulz, Fischbacher, Thöni, & Utikal, 2014).

The SHH may help to explain why individuals from certain societies are more likely to be cooperative and trusting than those from others. Although mechanisms such as direct and indirect reciprocity are likely to be fairly universal across cultures, the quality of institutions varies dramatically from society to society (Ellingsen, Herrmann, Nowak, Rand, & Tarnita, 2012; Gächter & Herrmann, 2009; Henrich et al., 2010; Herrmann, Thoni, & Gächter, 2008). This variation creates environments where it is safer versus more dangerous to trust strangers. Thus, it may be that people living under effective institutions internalize cooperative norms to a greater extent than those living under poorly functioning institutions.

It is difficult to draw strong causal conclusions about cross-cultural variation in cooperation. However, correlational evidence suggests that

institutions such as laws and markets play an important role in fostering cooperative norms. For example, people living in countries with less corruption and stronger rule of law were less likely to engage in antisocial punishment behavior in PGG games (Ellingsen et al., 2012; Gächter & Herrmann, 2009; Herrmann et al., 2008); furthermore, across a range of societies from hunter-gatherers to industrialized nations, greater market integration was associated with greater generosity in economic games (Henrich et al., 2010).

Direct experimental evidence also exists for the effect of social environment on heuristic-based cooperation in the form of experiments using repeated PDs as a model of the future consequences created by institutions. In these studies subjects were randomly assigned to play a series of repeated PDs either under a set of “good” rules that favored cooperation (large repeat probability) or a set of “bad” rules that favored defection (small repeat probability); afterward, all subjects played an identical battery of one-shot anonymous games (Peysakhovich & Rand, 2013). Subjects assigned to “good” environments were more cooperative, altruistic, trusting, and trustworthy in the subsequent one-shot games as well as more inclined to punish selfishness. Furthermore, this effect was especially strong among individuals who relied more on heuristic processing. These results provide direct evidence for spillover effects: the rules governing your interactions with others influence the cooperative heuristics you adopt.

## Conclusion

Cooperation with unrelated individuals is a hallmark of humankind despite the temptation to behave selfishly. Mechanisms that create future consequences for present actions can make cooperation pay off in the longer term and allow cooperation to arise and be maintained. Some of the most powerful of these mechanisms are direct reciprocity, indirect reciprocity, and institutions. The cooperative behavior promoted by these mechanisms is then implemented with the help of various proximate psychological processes such as emotions like empathy. We argue that social heuristics form the bridge between ultimate and proximate explanations for cooperation: we internalize advantageous behavioral strategies as intuitive defaults, and thus mechanisms that promote cooperation may lead us to cooperate automatically, even in specific situations that are beyond the reach of

any mechanism. Such “irrational” cooperation is central to the success of human civilization and forms a key component of morality.

We conclude this chapter by describing important open questions regarding human cooperative behavior. In this chapter we have proposed that intrinsic motivations for cooperative behavior arise as a result of extrinsic incentives that make cooperation advantageous. The evidence that we present for this hypothesis, however, runs counter to work showing that extrinsic incentives can *crowd out*, or undermine, intrinsic motivation (Titmuss, 1970). For example, in a classic study of extrinsic undermining, young children became less interested in drawing with markers when they were asked to draw in exchange for a “good player award” (Lepper, Greene, & Nisbett, 1973). Extrinsic undermining has also been shown to occur in social dilemma situations (e.g., Frey & Oberholzer-Gee, 1997; Gneezy & Rustichini, 2000). How can we reconcile these seemingly contradictory results? The answer likely concerns the *nature* of the incentives. For example, it may be that explicit financial incentives *crowd out* intrinsic motivation, whereas more implicit reputational incentives *crowd in* intrinsic motivation. Developing an understanding of how to apply incentives in a way that effectively creates cultures of cooperation with internalized prosocial norms is of great scientific interest as well as practical importance.

Another open question concerns how people *respond* to cooperative and selfish behaviors with behaviors such as reciprocity and norm enforcement. Not only do people engage in first-hand cooperative behaviors, but they also attend to and monitor the cooperative behaviors of others. Individuals respond to cooperative (and noncooperative) behavior both when the behavior has directly affected them (as *second parties*) and when it has not (as *third parties*). Evidence suggests that both second and third parties frequently reward cooperators and punish defectors, even at personal costs (Almenberg, Dreber, Apicella, & Rand, 2011; Fehr & Fischbacher, 2003, 2004). Reciprocity and norm enforcement are important behaviors to understand, as they create additional consequences for violating cooperative norms and thus help maintain cooperative societies.

As with cooperative behavior, the question arises as to why individuals engage in sanctioning behavior. Evidence suggests that second-party punishment is an intuitive and automatic response to unfair treatment (Grimm & Mengel, 2011; Rand & Nowak, 2013; Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003). This suggests that it may represent a generally

advantageous strategy—perhaps because it serves to deter future exploitation. But why would a third-party observer pay costs to punish selfish behavior when she has not been personally harmed? An understanding of the ultimate and proximate mechanisms that support third-party norm enforcement is needed. Research investigating these mechanisms will both contribute to our theoretical understanding of human cooperation and help us to understand how to best foster cultures of cooperation.

## References

- Almenberg, J., Dreber, A., Apicella, C. L., & Rand, D. G. (2011). *Third party reward and punishment: Group size, efficiency and public goods psychology and punishment*. Hauppauge, NY: Nova Science Publishers.
- Andreoni, J., & Gee, L. (2012). Gun for hire: Delegated enforcement and peer punishment in public goods provision. *Journal of Public Economics*, 96(11), 1036-1046.
- Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.
- Baldassari, D., & Grossman, G. (2011). Centralized sanctioning and legitimate authority promote cooperation in humans. *Proceedings of the National Academy of Sciences USA*, 108(27), 11023–11027.
- Barclay, P., & Willer, R. (2007). Partner choice creates competitive altruism in humans. *Proceedings. Biological Sciences*, 274(1610), 749–753.
- Batson, C. D., & Ahmad, N. (2001). Empathy induced altruism in a prisoner's dilemma II: What if the target of empathy has defected? *European Journal of Social Psychology*, 31(1), 25–36.
- Batson, C. D., Ahmad, N., & Lishner, D. A. (2011). Empathy and altruism. In *The Oxford handbook of positive psychology*, 417-427. Oxford: Oxford University Press.
- Batson, C. D., Duncan, B. D., Ackerman, P., Buckley, T., & Birch, K. (1981). Is empathic emotion a source of altruistic motivation? *Journal of Personality and Social Psychology*, 40(2), 290.
- Batson, C. D., Dyck, J. L., Brandt, J. R., Batson, J. G., Powell, A. L., McMaster, M. R., et al. (1988). Five studies testing two new egoistic alternatives to the empathy-altruism hypothesis. *Journal of Personality and Social Psychology*, 55(1), 52.
- Blonski, M., Ockenfels, P., & Spagnolo, G. (2011). Equilibrium selection in the repeated prisoner's dilemma: Axiomatic approach and experimental evidence. *American Economic Journal: Microeconomics*, 3(3), 164–192.
- Bowles, S., Choi, J., & Hopfensitz, A. (2003). The co-evolution of individual behaviors and social institutions. *Journal of Theoretical Biology*, 223(2), 135–147.

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

Cornelissen, G., Dewitte, S., & Warlop, L. (2011). Are social value orientations expressed automatically? Decision making in the dictator game. *Personality and Social Psychology Bulletin*, 37(8), 1080–1090.

Dal Bó, P., & Fréchet, G. R. (2011). The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review*, 101(1), 411–429.

Dreber, A., Fudenberg, D., & Rand, D. G. (2014). Who cooperates in repeated games? *Journal of Economic Behavior & Organization*, 98, 41–55.

Dreber, A., Rand, D. G., Fudenberg, D., & Nowak, M. A. (2008). Winners don't punish. *Nature*, 452(7185), 348–351.

Ellingsen, T., Herrmann, B., Nowak, M. A., Rand, D. G., & Tarnita, C. E. (2012). Civic capital in two cultures: The nature of cooperation in Romania and USA. Available at SSRN: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2179575](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2179575).

Engel, C., & Rand, D. G. (2014). What does “clean” really mean? The implicit framing of decontextualized experiments. *Economics Letters*, 122(3), 386–389.

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, 425(6960), 785–791.

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, 25(2), 63–87.

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), 1015–1030.

Frey, B. S., & Oberholzer-Gee, F. (1997). The cost of price incentives: An empirical analysis of motivation crowding-out. *American Economic Review*, 87(4), 746–755.

Fudenberg, D., & Maskin, E. S. (1986). The Folk Theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54(3), 533–554.

Fudenberg, D., Rand, D. G., & Dreber, A. (2012). Slow to anger and fast to forgive: Cooperation in an uncertain world. *American Economic Review*, 102(2), 720–749.

Gächter, S., & Herrmann, B. (2009). Reciprocity, culture and human cooperation: Previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 364(1518), 791–806.

Gneezy, U., & Rustichini, A. (2000). Pay enough or don't pay at all. *Quarterly Journal of Economics*, 115(3), 791–810.

Greif, A. (1993). Contract enforceability and economic institutions in early trade: The Maghribi traders' coalition. *American Economic Review*, 83(3), 525–548.

Grimm, V., & Mengel, F. (2011). Let me sleep on it: Delay reduces rejection rates in ultimatum games. *Economics Letters*, *111*(2), 113–115.

Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior*, *26*, 245–256.

Hauge, K. E., Brekke, K. A., Johansson, L.-O., Johansson-Stenman, O., & Svedsäter, H. (2009). *Are social preferences skin deep? Dictators under cognitive load*. University of Gothenburg Working Papers in Economics, (371).

Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., et al. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science*, *327*(5972), 1480–1484.

Herrmann, B., Thoni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science*, *319*(5868), 1362–1367.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*(9), 697–720.

Karlan, D., & McConnell, M. A. (2012). *Hey look at me: The effect of giving circles on giving*. National Bureau of Economic Research, Cambridge, MA.

Lacetera, N., & Macis, M. (2010). Social image concerns and prosocial behavior: Field evidence from a nonlinear incentive scheme. *Journal of Economic Behavior & Organization*, *76*(2), 225–237.

Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, *28*(1), 129.

Milgrom, P. R., & North, D. C. (1990). The role of institutions in the revival of trade: The law merchant, private judges, and the champagne fairs. *Economics and Politics*, *2*(1), 1–23.

Milinski, M., Semmann, D., Bakker, T. C. M., & Krambeck, H. Jr. (2001). Cooperation through indirect reciprocity: Image scoring or standing strategy? *Proceedings of the Royal Society of London. B, Biological Sciences*, *268*(1484), 2495–2501.

North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge, UK: Cambridge University Press.

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, *314*(5805), 1560–1563.

Nowak, M. A., & Sigmund, K. (1992). Tit for tat in heterogeneous populations. *Nature*, *355*, 250–253.

- Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, 437(7063), 1291–1298.
- Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge, UK: Cambridge University Press.
- Ouss, A., & Peysakhovich, A. (2013). When punishment doesn't pay: "Cold Glow" and decisions to punish. Available at SSRN: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2247446](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2247446).
- Peysakhovich, A., & Rand, D. G. (2013). Habits of virtue: Creating norms of cooperation and defection in the laboratory. Available at SSRN: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2294242](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2294242).
- Pfeiffer, T., Tran, L., Krumme, C., & Rand, D. G. (2012). The value of reputation. *Journal of the Royal Society, Interface*, 9(76), 2791–2797.
- Rand, D. G., Fudenberg, D., & Dreber, A. (2013). It's the thought that counts: The role of intentions in reciprocal altruism. Available at SSRN: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2259407](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2259407).
- Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, 489(7416), 427–430.
- Rand, D. G., & Kraft-Todd, G. T. (in press). Reflection does not undermine self-interested prosociality. *Frontiers in Behavioral Neuroscience*.
- Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in Cognitive Sciences*, 17(8), 413–425.
- Rand, D. G., Ohtsuki, H., & Nowak, M. A. (2009). Direct reciprocity with costly punishment: Generous tit-for-tat prevails. *Journal of Theoretical Biology*, 256(1), 45–57.
- Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., et al. (2014). Social heuristics shape intuitive cooperation. *Nature Communications*, 5, 3677.
- Roch, S. G., Lane, J. A. S., Samuelson, C. D., Allison, S. T., & Dent, J. L. (2000). Cognitive load and the equality heuristic: A two-stage model of resource overconsumption in small groups. *Organizational Behavior and Human Decision Processes*, 83(2), 185–212.
- Ruff, C. C., Ugazio, G., & Fehr, E. (2013). Changing social norm compliance with noninvasive brain stimulation. *Science*, 342, 482–484.
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, 300(5626), 1755–1758.

Schulz, J. F., Fischbacher, U., Thöni, C., & Utikal, V. (2014). Affect and fairness: Dictator games under cognitive load. *Journal of Economic Psychology* 41, 77–87.

Slooman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119(1), 3.

Sparks, A., & Barclay, P. (2013). Eye images increase generosity, but not for long: The limited effect of a false cue. *Evolution and Human Behavior*, 34(5), 317–322.

Tinghög, G., Andersson, D., Bonn, C., Böttiger, H., Josephson, C., Lundgren, G., et al. (2013). Intuition and cooperation reconsidered. *Nature*, 497(7452), E1–E2.

Titmuss, R. (1970). *The gift relationships*. London: George Allen.

Wedekind, C., & Milinski, M. (2000). Cooperation through image scoring in humans. *Science*, 288(5467), 850–852.

Yoeli, E., Hoffman, M., Rand, D. G., & Nowak, M. A. (2013). Powering up with indirect reciprocity in a large-scale field experiment. *Proceedings of the National Academy of Sciences USA*, 110(Supplement 2), 10424–10429.

