Harnessing the Power of Reputation: Strengths and Limits for Promoting Cooperative Behaviors

Pat Barclay, Department of Psychology, University of Guelph, Guelph, ON, Canada. Email: barclayp@uoguelph.ca.

Abstract: Evolutionary approaches have done much to identify the pressures that select for cooperative sentiment. This helps us understand when and why cooperation will arise, and applied research shows how these pressures can be harnessed to promote various types of cooperation. In particular, recent evidence shows how opportunities to acquire a good reputation can promote cooperation in laboratory and applied settings. Cooperation can be promoted by tapping into forces like indirect reciprocity, costly signaling, and competitive altruism. When individuals help others, they receive reputational benefits (or avoid reputational costs), and this gives people an incentive to help. Such findings can be applied to promote many kinds of helping and cooperation, including charitable donations, tax compliance, sustainable and pro-environmental behaviors, risky heroism, and more. Despite the potential advantages of using reputation to promote positive behaviors, there are several risks and limits. Under some circumstances, opportunities for reputation will be ineffective or promote harmful behaviors. By better understanding the dynamics of reputation and the circumstances under which cooperation can evolve, we can better design social systems to increase the rate of cooperation and reduce conflict.

Keywords: cooperation, altruism, helping, applied game theory, costly signaling, reciprocity

Why Help?

Human cooperation is truly impressive. Humans regularly incur great costs to help kin and non-kin. In terms of formal helping, in Canada in 2007 alone, 46% of people volunteered 166 hours, on average, and 84% of people donated a total of almost 10 billion dollars (Hall, Lasby, Ayer, and Gibbons, 2009). This formal help is dwarfed by the vast amount of informal helping performed in favors, information exchange, tips, and even restraint from performing selfish behaviors. Canadians are by no means alone in their generosity, so we must ask: why are humans so generous to others, even towards non-kin?
A first pass glance at natural selection would suggest that helping others would be disfavored when there are costs associated with helping, because those who avoid the costs would have an advantage over those who pay them. Evolutionists have long known that organisms do not act “for the good of the species”. Nevertheless, costly helping is ubiquitous, so the question is: why? What factors have caused cooperative sentiments to arise? This is a central puzzle in evolution. An answer to this question can help us to foster helpful behaviors by creating situations to promote helping. Given the importance of cooperation on both local and global scales, it is crucial to understand the selective pressures that maintain all forms of cooperation.

Why a Functional Approach is Necessary for Promoting Cooperation

Much research seeks to identify the proximate psychological mechanisms that trigger helping behavior within individuals (e.g., neural pathways, specific emotions like empathy (Batson et al., 1997) or “warm glow” (Andreoni, 1990)). Other research examines how these psychological mechanisms develop within an individual’s lifetime (e.g., learning (Clark, 1975), gene-environment interactions (Knafo and Plomin, 2006)). These are both important questions, but it is also important to understand the function of these cooperative sentiments. Why do they exist at all? What are the advantages of possessing them? And what selective pressures cause them to persist? Together, these three research areas – mechanism, development, and evolutionary function – represent three of Tinbergen’s (1968) four “Levels of Analysis” (the fourth being phylogeny), and an understanding of all four levels is necessary.

In particular, it is important to understand evolutionary function in order to produce social change. By understanding the function of cooperative behavior, we can change social situations to allow for more of that behavior. For example, which would be more effective: a) understanding what selective pressures have caused cooperative sentiment to arise, and then using that knowledge to create (or mimic) situations that are conducive to long-term cooperation because those selective forces are present; or b) investigating what specific emotions are involved in charity in order to trigger those emotions, hoping that the triggered emotional response is specific enough to affect charitable behavior, and hoping that it is not subject to habituation or extinction from a lack of reinforcement? Compared with the former, the latter seems hopelessly indirect, temporary, and possibly ineffective. This is why identifying the function of cooperative sentiment is paramount for producing lasting change.

To understand function, we need to know what consequences a cooperative person tends to experience. Our evolved prosocial sentiments cause us to perform helpful behavior, so how do others respond to this? I argue that cooperative behavior often produces reputational benefits, which outweigh the costs. However, it is important to note that not every instance of helping will result in such benefits, nor does it need to. Our evolved sentiments may occasionally be triggered in ways that do not advance our fitness (Haley and Fessler, 2005; West, El Mouden, and Gardner, 2011), and occasional errors are inevitable in any decision-making process (e.g., Nesse, 2005; Todd, 2001). Instead, prosocial sentiments like empathy are adaptive on average, because reputational
opportunities are often present and important, and the average reputational benefits of acting on those sentiments outweigh the costs of occasionally helping when one cannot acquire a reputation (Delton, Krasnow, Cosmides, and Tooby, 2012). As such, reputation can continue to have a powerful effect on helpful behavior as long as real reputational consequences arise frequently enough to make such helping occasionally worthwhile.

**Why Generalized Helping?**

There is much research on helping that benefits kin (e.g., Hamilton, 1964), has direct personal benefits to the helper (e.g., Diekmann, 1985, Kümmerli et al., 2007), benefits someone whom the helper has some stake in (Kokko, Johnstone, and Clutton-Brock et al., 2001; Roberts, 2005), or is directly reciprocated by the recipient (Axelrod, 1984; Trivers, 1971). These situations are well-studied in evolutionary studies, and I have reviewed them and their applications elsewhere (Barclay, 2011a; Barclay and Van Vugt, in press). Because humans also cooperate with those who are neither kin nor reciprocal partners, I will focus on one specific and particularly selective pressure – reputation – in its various forms. By understanding these selective pressures, we can learn to harness them to increase helpful behaviors.

**Indirect Reciprocity**

People who help others tend to receive help when they need it. Some of this help is directly reciprocated by the recipient in the manner of “you scratch my back and I’ll scratch yours” (i.e., “direct reciprocity”) (e.g., Axelrod, 1984). In addition, help is often indirectly reciprocated by people other than the recipient (i.e., “indirect reciprocity”) (Nowak and Sigmund, 2005). Helpers earn a good reputation, and this makes them more likely to receive aid than non-helpers. This pattern is shown to be evolutionarily stable in many mathematical models and game theoretical simulations (reviewed by Nowak and Sigmund, 2005), and many experimental and field studies have shown that those who give are more likely to receive, even from those whom they have never personally given to (e.g., Bolton, Katok, and Ockenfels, 2005; Gurven, Allen-Arave, Hill, and Hurtado, 2000; Milinski, Semmann, Bakker, and Krambeck, 2001; Seinen and Schram, 2006; Wedekind and Milinski, 2000). There is even evidence for indirect reciprocity in cleaner fish: client fish avoid cleaners who are observed to give “bad service” to other clients; good cleaners receive more clients (Bshary and Grutter, 2006).

**Punishment**

Those who refuse to cooperate do not only face a lack of rewards; they may also experience active punishment from others. Such punishment can include monetary fines (Fehr and Gächter, 2002; Ostrom, 1990), public shaming and criticism (Barr, 2001; Fessler, 2002), negative gossip (Piazza and Bering, 2008), and possibly physical threats. These punishments promote cooperation because they make it costly to refuse to cooperate. In a recent example, the British government shamed comedian Jimmy Carr into stopping a dubious tax avoidance scheme; this reputation-based shaming had strong tangible consequences because it exposed actions that risked alienating his audiences (thereby
Laboratory and field data show that people are very willing to impose costs upon others who do not contribute their fair share of time, effort, or money to a common cause (e.g., Cordell and McKean, 1992; Ostrom, Walker, and Gardner, 1992; Price, 2005; Yamagishi, 1986). Studies using brain-scanning technology suggest that people experience pleasure when non-cooperators receive punishment (de Quervain et al., 2004; Singer et al., 2006). Punishment and indirect reciprocity can work together to promote cooperation (Rockenbach and Milinski, 2006), though punishment may be more cost-effective in situations where unanimous cooperation is required (Oliver, 1980). But why be the one to perform the punishment, given the personal costs in terms of time, effort, and retaliation risk (Oliver, 1980; Yamagishi, 1986)? One solution is that punishers receive reputational benefits by signaling their dislike of unfairness, ability to punish, and intolerance of being exploited (Barclay, 2006, submitted; Horita, 2010; Nelissen, 2008; though see Kiyonari and Barclay, 2008). Altogether, punishment supports cooperation by targeting those with bad reputations, and reputation can support punishment itself.

Costly Signaling

Suppose that you possess a desirable quality that you want to advertise to others, such as wealth, genetic quality, or good character. How can you do so believably? One way is to perform or possess something that would be too costly to be worth it for someone who did not honestly possess that quality (Searcy and Nowicki, 2005). For example, a peacock’s tail is very costly to build, and the benefits of possessing a big one are only worth it for a high quality peacock (Zahavi and Zahavi, 1997). Similarly, Ferraris and mansions likely function as an honest display of wealth in humans (Veblen, 1899/1994). Such signaling need not be conscious, just as a peacock does not consciously grow its tail to attract mates.

Extravagant generosity may similarly function as a (probably unconscious) display of abilities or resource-holding potential. For example, philanthropy requires money to spare, and Bill Gates and Warren Buffett are known in North America as much for their philanthropy as for their wealth. Hunting big game requires many abilities, and good hunters receive widespread recognition when they share their kills (e.g., Smith and Bliege Bird, 2000); as a result, they have higher reproductive success than non-hunters (Smith, 2004), including more extra-marital affairs (Kaplan and Hill, 1985). Heroes are praised for successfully taking physical risks for others, and recent evidence suggests they are more desirable as mates (Farthing, 2005; Kelly and Dunbar, 2001). Audiences attend and respond to such acts because they carry useful information about the helper’s abilities and/or resources, and the possibility of receiving something makes such acts particularly attractive to attend to (Smith and Bliege Bird, 2000).

In addition to signaling resources or abilities, helpful behavior can signal cooperative intent and concern for others. In long-term interactions, helpful behavior at time A would not be worth it for someone who intends to defect at time B and face mutual defection in each subsequent interaction (André, 2010). As such, one’s helping at time A predicts one’s helping at time B, albeit imperfectly (e.g., Kurzban and Houser, 2005). Not surprisingly, people are more trusting of those who give money to charities or to fellow group members (e.g., Albert, Güth, Kirchler, and Maciejovsky, 2007; Barclay, 2004, 2006;
Harnessing the power of reputation

Keser, 2003). The costs of religious rituals may also function as signals of cooperative intent because the subjective costs would not be worth it for non-believers. As a result, costly religious rituals are associated with higher levels of trust (Sosis, 2004). Charitable donations are often religious donations (Hall et al., 2009). It is also possible that other groups require costly rituals, volunteering, or philanthropy in order to screen out uncooperative applicants (e.g., fraternities, Masons, medical schools). For a more thorough review of cooperation as an honest signal of abilities, resources, and cooperative intent, and the empirical evidence for each, see Barclay (2010b).

Harnessing the Power of Reputation

The three theories above have their differences, but they all rely on people acquiring reputations for (non-)cooperation. As such, their most obvious prediction is that people will be more cooperative when they can acquire a good reputation (e.g., if their cooperation gets announced to others, or if they receive some badge indicative of their cooperation). This prediction has been abundantly confirmed in many domains, including experimental games (e.g., Barclay, 2004; Hardy and Van Vugt, 2006; Hoffman, McCabe, Schachat, and Smith, 1994; Milinski, Semmann, and Krambeck, 2002; Rege and Telle, 2004), volunteering (Bereczkei, Birka, and Kerekes, 2007), contributions towards educating others about climate change (Milinski, Semmann, Krambeck, and Marotzke, 2006), charitable donations (Harbaugh, 1998; Karlan and McConnell, 2012), voter turnout (Gerber, Green, and Larimer, 2008), blood donations (Low and Heinen, 1993), socially responsible purchasing (Kimura et al., 2012), and simulated tax donations (Coricelli, Joffily, Montmarquette, and Villeval, 2010). Making people think about status can increase benevolence (Griskevicius et al., 2007). Even fear of gossip or disapproval can stimulate cooperation (Masclet, Noussair, Tucker, and Villeval, 2003; Piazza and Bering, 2008; Xiao and Houser, 2009). This increased cooperation is sometimes strategic (Semmann, Krambeck, and Milinski, 2004; Barclay and Willer, 2007), but not always; observation might even cause a genuine increase in cooperative sentiments such as empathy, shame, or guilt. Whether strategic or genuine, it should be easy to add reputational opportunities to many systems, which will increase people’s cooperation (e.g., announcing people’s cooperation in any domain, or giving identifiable cues or badges to cooperators) (Griskevicius, Cantú, and Van Vugt, 2012; Low and Heinen, 1993). For an additional review of this work, see Griskevicius and colleagues (2012).

People are so sensitive to their reputations that they will respond to minimal cues of being observed. Haley and Fessler (2005) showed that people donated more money to their partners in an experimental game when there were stylized eyespots on their computer screen than when such eyespots were absent. These results have been replicated numerous times with other types of eyes (Burnham and Hare, 2007; Mifune, Hashimoto, and Yamagishi, 2010) including three dots in an eyes-like configuration (Rigdon, Ishii, Watabe, and Kitayama, 2009), and seem to be driven by implicit expectations of a good reputation (Oda, Niwa, Honma, and Hiraishi, 2011). Field experiments have shown that the presence of eyes increases the amounts that people pay for coffee on an “honour system” (Bateson et al., 2006) and reduces the amount of litter that people leave in cafeterias (Ernest-Jones,
Harnessing the power of reputation

Nettle, and Bateson, 2011). This all suggests that reputation can be harnessed even with minimal cues, though the effects of such minimal cues may only be transient (see limitation #3 below; Sparks and Barclay, submitted).

People give more when observed, and they give the most when competing over partnerships (Barclay and Willer, 2007). Engaging in competitive altruism makes people more desirable as partners and gives them greater access to social partnerships (Barclay, 2004; Barclay and Willer, 2007; Roberts, 1998; Sylwester and Roberts, 2010), sometimes including romantic partnerships (Barclay, 2010a; Iredale, Van Vugt, and Dunbar, 2008). In such a market for partners, it pays to “keep up with the Joneses’ generosity” and possibly exceed it in order to get access to the best partners. Generosity is a relative term: someone who gives more than average looks good, whereas someone who gives less than average looks bad and risks losing the competition for partners. As such, people are highly sensitive to others’ cooperation, for example, they give more money to radio fundraising when told about others’ high contributions (Shang and Croson, 2006), and make more pro-environmental decisions like cutting energy usage or re-using hotel towels when told that most others do so (Allcott, 2011; Goldstein, Griskevicius, and Cialdini, 2007). Competitive altruism explains why people are so sensitive about appearing less generous, as well as why generosity can sometimes escalate. People are naturally competitive, and we can harness this to foster competitive altruism including competitive sustainability (Griskevicius et al., 2012). For example, we can advertise the contributions of the best cooperators, publicly name the biggest philanthropists or the most prosocial individuals and corporations (Duffy and Kornienko, 2006; Griskevicius et al., 2012), and increase people’s ability to meet high cooperators and assort into the most cooperative groups and partnerships (Barclay, 2011b).

Reputations are also important in business and consumerism (Frank, 2004). Many companies and consumers are aware of this, and trader reputations are already used to facilitate trust on Internet sites such as eBay, Amazon, and TripAdvisor. Consumers may engage in “conspicuous conservation” when buying “green” products like a Toyota Prius (Sexton and Sexton, 2011), especially after they are primed to think about their own reputations (Griskevicius, Tybur, and Van den Bergh, 2010), or when they are presented with social comparisons about what others do (Allcott, 2011; Goldstein et al., 2007). For example, people purchase more fair trade products when being observed (Kimura et al., 2012). Because of this, we can harness reputations to promote socially responsible consumerism (or at least less harmful consumerism). This can affect corporate behavior; companies often care about their reputation and this indicates that they are aware of the need to appear somewhat socially responsible (or at least not harmful). After all, corporations respond to market demands, and consumer boycotts have had some success at influencing corporations. Having more competitors will increase companies’ need for social responsibility (or appearance thereof), as market size affects how generous one needs to be in social competition (Barclay, 2011b). Furthermore, individual CEOs can be made to care about their personal reputations if they are seen as responsible for the socially harmful or beneficial policies of their companies. Thus, attaching reputations to socially (ir)responsible products can alter both consumption and production patterns.
When Will Reputation Work?

The effectiveness of reputation at promoting an action will depend on how that action is viewed. If an act is seen as “good”, then reputational opportunities will result in people performing it more often. In particular, “good” people will perform it more often when the costs deter “bad” people, such that the act carries useful information about a person’s character. Also, “good” people should not do “bad things” knowingly (though see limitation #7 below).

What determines whether an act is seen as “good”, “bad”, or “neutral”? This question is not always easy to answer. Some acts are almost intrinsically good, such as saving others’ lives. Other acts are potentially ambiguous: is public smoking good, bad, or neutral? What about drunk driving? Recycling? Purchasing biofuels? Supporting Cuba? One must have information on how these acts help or harm others, otherwise they may appear neutral. Someone who knows the harms (or benefits) of an act but does (or does not do) it anyway can be inferred to be a “bad” person. Drinking and driving campaigns have been successful at broadcasting these harms and creating reputational pressures against drunk driving, and campaigns against littering and second-hand smoke have also enjoyed some success. British propaganda in World War II generated social norms about conservation. One can imagine similar campaigns to increase awareness of the social harms of pollution, unsustainable consumption, and buying socially irresponsible products, and to shame those who take part in these activities. Similarly, one could educate about the shared benefits of sustainability, recycling, public transit usage, full tax compliance, tree-planting, voting, social action, or any other campaign (see also Kazdin, 2009).

This education effectively increases reputational pressures. For example, people are particularly likely to donate to the fight against climate change when they are non-anonymous if they have already been educated about the dangers of climate change (Milinski, van Dijk, DeCremer, and Wilke, 2006). Thus, education and reputation work together such that education creates the reputational pressures, and reputational pressures give people a big reason to care. Without reputational pressures, cooperation can decline even among people who are fully aware of the effects of their actions - this is the logic of the “tragedy of the commons” that reputation “solves” (Hardin, 1968; Milinski et al., 2002).

Limitations and Unknowns about Harnessing Reputation

There are currently several important limitations and unknowns associated with harnessing reputation, all of which could undermine or reverse any attempts to promote positive behavior. The following limitations are in approximately increasing order of how problematic each one is, and includes potential remedies (see also Barclay, 2011a).

1) Reputational benefits must outweigh the cost of helping

To be effective, incentives must be sufficient to outweigh costs of acting. This point, although obvious, is easy to forget in practice. If the expected reputational benefits do not outweigh the costs of helping, then they will not be effective at eliciting helping, and
may even suppress helping (see “crowding out” in limitation #6).

2) Reputational cues must be stronger than other situational factors

Some cues of reputation are very strong, like immediate monetary consequences or a large audience of attractive members of the opposite sex. Other cues are relatively weak, like eyespots in the background of one’s environment (Haley and Fessler, 2005). Naturally, stronger situations or cues should have more impact than weaker ones (e.g., Xiao and Houser, 2009; see also Cooper and Withey, 2009). When non-reputational factors are present and strong, like strong internalized social norms in a situation, then these may overwhelm weak reputational cues. More research should determine which cues (e.g., sex, status, and group membership of audiences) are most important to reputation.

3) People may habituate to non-informative cues of reputation

Humans are excellent learners, which includes eventually ceasing behaviors that receive no reinforcement. If cues of reputation are not eventually followed by actual reputation that matters, then people should eventually come to ignore these cues. For example, eyespots are a (false) cue of reputation, and Sparks and Barclay (submitted) found that eyespots lose their effectiveness at eliciting cooperation in experimental games if the eyes are constantly present. Soetevent (2005) found that reputational cues were only temporarily effective in eliciting church donations. We might also predict that people will eventually respond less to the scrutiny of audiences who do not react to cooperation or free-riding, or who cannot have any tangible effect on one’s status or fitness (including by communicating with fitness-relevant audiences). To remain effective, false cues of reputation (e.g., eyespots) must be changed frequently, or occasionally followed by real observation, and observation must occasionally be followed by reaction.

4) Not everyone values reputation

Some people may not need or value reputation (Barclay and Reeve, 2012; Rege and Telle, 2004), such that reputational incentives will not work on them. These can include individualists, high status people, autistic people, or psychopaths. These types may require obvious, tangible, immediate, and strong reward or punishment. Alternately, social systems should be designed to avoid having these non-responders drag down the cooperation of others, possibly by focusing on the most cooperative people rather than the least cooperative.

5) Reputation only pays off in the long-term

It is only worth acquiring a reputation if one will be around long enough to eventually benefit from it. Accordingly, people who devalue the future tend to cooperate less in experimental games (Harris and Madden, 2002), and are predicted to value sustainability less (Griskvicius et al., 2012). Luckily, people’s valuation of the future can vary adaptively according to situational cues (Wilson and Daly, 2004) and life circumstances (reviewed by Griskevicius et al., 2012). As such, anything that increases a person’s perceived “shadow of the future” should increase their sensitivity to reputation, including a perceived future within a group.
6) **Extrinsic incentives “crowd out” intrinsic motivations**

Many people seem to have internalized norms about helping others such that they genuinely value helping (Fehr and Fischbacher, 2004; Simpson and Willer, 2008). However, much research shows that extrinsic incentives can reduce people’s intrinsic motivations (for reviews, see Deci, Koestner, and Ryan, 1999; Bowles, 2008). In addition to being ineffective, extrinsic incentives can make situations even worse: Gneezy and Rustichini (2000) found that parents were more likely to be late to pick up their children from daycare when there were fines present. Similarly, punishment of non-cooperators undermines trust, possibly because it obscures people’s motivations for cooperating (Mulder et al., 2006). Because of this, would-be social engineers must be very careful that any extrinsic reputational benefits do not reduce or “crowd out” intrinsic motivations for cooperation. One solution is to use implicit incentives like verbal rewards, which do not reduce intrinsic motivations (Deci et al., 1999), yet can still be effective at promoting positive behaviors (Barr, 2001; Masclet et al., 2003). Another solution is to give people opportunities to create their own informal sanctions rather than relying on top-down explicit incentives (though see limitation #4 above).

7) **Reputation can promote negative behaviors**

Wealth and power are often signaled through non-generous behaviors, such as conspicuous consumption (Frank, 2000; Veblen, 1899/1994) or physical and social displays which do not benefit others. Furthermore, individuals may also need to establish their toughness, unwillingness to tolerate exploitation, or risk-acceptance (e.g., Chagnon, 1997; Daly and Wilson, 1988). As such, under some circumstances, reputational opportunities can increase negative behaviors such as aggression (Griskevicius et al., 2009) and risk-taking (Chen, Baker, Braver, and Li, 2000). In environments where such traits are valued, would-be social engineers must ensure that they only create reputational opportunities for positive behaviors (e.g., broadcasting people’s cooperation but not their other behaviors) or they must allow negative behaviors to be channeled toward good ends (e.g., signaling toughness by punishing non-cooperators: Barclay, submitted).

8) **Reputations can be manipulated**

People will sometimes give to others when observed, only to be selfish when anonymous (Barclay and Willer, 2007; Semmann et al., 2004). Gossip can also be used to falsely increase or decrease someone’s reputation (Hess and Hagen, 2006). Corporations are well aware of this: “green-washing” refers to attempts to make environmentally harmful products or actions appear environmentally responsible, and companies may try to manipulate their online reputations on internet sites like eBay, Amazon, or TripAdvisor. The greater the reputational pressures, the more worthwhile it is to manipulate one’s reputation or the reputation of one’s competitors, even at significant cost. Potential solutions to this include: increasing reputations to include all information necessary to judge an issue, using public costly signals to guarantee honesty and commitment, or increasing the degree of partner (or consumer) choice to make it easier to replace a potentially dishonest partner with a genuine one. On a more public scale, increased funding for watchdogs such as “Truth in Advertising” would help maintain the accuracy of the
information that people use to make their consumer choices.

9) Identifying reputational incentives may reduce them

If you thought that someone was acting morally solely to receive social benefits, would you still trust them as much? Probably not. When someone’s motivation for helping is unclear, it can reduce the reputational benefits that he or she receives because it makes the helping less effective as a signal of good character. People may even deliberately call attention to such benefits in order to downplay the generosity of their competitors. This may be currently difficult to resolve because there is little theoretical or empirical work on the details of how signals of cooperative intent work. Some potential solutions include: a) creating subtle opportunities for reputational benefit while strictly avoiding mention of such benefits; b) announcing people’s generosity on their behalf so they can avoid “shameless self-promotion”; c) comparing people who acquire a reputation via helping with people who don’t help at all, or with people who acquire a reputation in less positive ways like conspicuous consumption. A focus on the appropriate comparisons should help maintain at least some reputational benefits.

Conclusions

Reputation represents a largely untapped resource for eliciting all types of cooperative behaviors. The above limitations and unknowns about reputation are not insurmountable problems, as careful planning can circumvent them. Some may inspire more theoretical and empirical research about the dynamics and creation of reputation. Because reputation focuses on the ultimate function(s) of cooperation instead of merely its proximate psychological causes, a mature science of reputation can help to produce long-lasting solutions for promoting cooperation.

Acknowledgements: This work was supported by the Social Sciences and Humanities Research Council (SSHRC).

Received 4 May 2012; Revision submitted 17 July 2012; Accepted 30 July 2012

References

Albert, M., Güth, W., Kirchler, E., and Maciejovsky, B. (2007). Are we nice(r) to nice(r) people? An experimental analysis. Experimental Economics, 10, 53-69.
Allcott, H. (2011). Social norms and energy conservation. Journal of Public Economics, 95, 1082-1095.
André, J. B. (2010). The evolution of reciprocity: Social types or social incentives? The American Naturalist, 175, 197-210.
Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. The Economic Journal, 100, 464-477.
Axelrod, R. (1984). The evolution of cooperation. New York: Basic Books.
Barclay, P. (2004). Trustworthiness and competitive altruism can also solve the “tragedy of
Harnessing the power of reputation

the commons.” Evolution and Human Behavior, 25, 209-220.
Barclay, P. (2006). Reputational benefits for altruistic punishment. Evolution and Human Behavior, 27, 325-344.
Barclay, P. (2010a). Altruism as a courtship display: Some effects of third-party generosity on audience perceptions. British Journal of Psychology, 101, 123-135.
Barclay, P. (2010b). Reputation and the evolution of generous behavior. Hauppauge, NY: Nova Science Publishers.
Barclay, P. (2011a). The evolution of charitable behaviour and the power of reputation. In S. C. Roberts (Ed.), Applied evolutionary psychology (pp. 149-172). Oxford, UK: Oxford University Press.
Barclay, P. (2011b). Competitive helping increases with the size of biological markets and invades defection. Journal of Theoretical Biology, 281, 47-55.
Barclay, P. (submitted). “Don’t mess with the enforcer”: Deterrence as an individual-level benefit for punishing free-riders. Submitted for publication.
Barclay, P., and Reeve, H. K. (2012). The varying relationship between helping and individual quality. Behavioral Ecology, 23, 693-698.
Barclay, P., and Van Vugt, M. (in press). The evolutionary psychology of human prosociality: Adaptations, mistakes, and byproducts. In D. Schroeder and W. Graziaono (Eds.), Oxford handbook of prosocial behavior. Oxford, UK: Oxford University Press.
Barclay, P., and Willer, R. (2007). Partner choices creates competitive altruism in humans. Proceedings of the Royal Society of London B, 274, 749-753.
Barr, A. (2001). Social dilemmas and shame-based sanctions: Experimental results from rural Zimbabwe (Working Paper No. WPS/2001.11). Centre for the Study of African Economics, University of Oxford, Oxford, U.K.
Bateson, M., Nettle, D., and Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. Biology Letters, 2, 412-414.
Batson, C. D., Sager, K., Garst, E., Kang, M., Rubchinsky, K., and Dawson, K. (1997). Is empathy-induced helping due to self-other merging? Journal of Personality and Social Psychology, 73, 495-509.
Bereczkei, T., Birkas, B., and Kerekes, Z. (2007). Public charity offer as a proximate factor of evolved reputation-building strategy: An experimental analysis of a real-life situation. Evolution and Human Behavior, 28, 277-284.
Bolton, G. E., Katok, E., and Ockenfels, A. (2005). Cooperation among strangers with limited information about reputation. Journal of Public Economics, 89, 1457-1468.
Bowles, S. (2008). Policies designed for self-interested citizens may undermine “the moral sentiments”: Evidence from economic experiments. Science, 320, 1605-1609.
Bshary, R., and Grutter, A. (2006). Image scoring and cooperation in a cleaner fish mutualism. Nature, 441, 975-978.
Burnham, T., and Hare, B. (2007). Engineering human cooperation: Does involuntary neural activation increase public goods contributions? Human Nature, 18, 88-108.
Chagnon, N. (1997). Yanomamö (5th ed.). Wadsworth.
Chen, L. H., Baker, S. P., Braver, E. R., and Li, G. (2000). Carrying passengers as a risk factor for crashes fatal to 16- and 17-year-old drivers. Journal of the American
Harnessing the power of reputation

Medical Association, 283, 1578-1582.

Clark, R. D. (1975). The effects of reinforcement, punishment and dependency on helping behavior. Personality and Social Psychology Bulletin, 1, 596-599.

Cooper, W. H., and Withey, M. J. (2009). The strong situation hypothesis. Personality and Social Psychology Review, 13, 62-72.

Cordell, J., and McKean, M. A. (1992). Sea tenure in Bahia, Brazil. In D. W. Bromley (Ed.), Making the Commons Work: Theory, Practice, and Policy (pp. 183-205). San Francisco: ICS Press.

Coricelli, G., Joffily, M., Montmarquette, C., and Villeval, M. C. (2010). Cheating, emotions, and rationality: An experiment on tax evasion. Experimental Economics, 13, 226-247.

Daly, M., and Wilson, M. (1988). Homicide. New York: Aldine de Gruyter.

de Quervain, D. J. F., Fischbacher, U., Treyer, V., Schellhammer, M., Schneyder, U., Buck, A., and Fehr, E. (2004). The neural basis of altruistic punishment. Science, 305, 1254-1258.

Deci, E. L., Koestner, R., and Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on internal motivation. Psychological Bulletin, 125, 627-668.

Delton, A. W., Krasnow, M. M., Cosmides, L., and Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. PNAS, 108, 13335-13340.

Diekmann, A. (1985). Volunteer’s dilemma. Journal of Conflict Resolution, 29, 605-610.

Duffy, J., and Kornienko, T. (2006). Does competition affect giving? (Working Paper No. 275). Department of Economics, University of Pittsburgh, Pittsburgh, PA.

Ernest-Jones, M., Nettle, D., and Bateson, M. (2011). Effects of eye images on everyday cooperative behavior: A field experiment. Evolution and Human Behavior, 32, 172-178.

Farthing, G. W. (2005). Attitudes toward heroic and nonheroic physical risk takers as mates and friends. Evolution and Human Behavior, 26, 171-185.

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

Fehr, E., and Fischbacher, U. (2004). Social norms and human cooperation. Trends in Cognitive Sciences, 8, 185-190.

Fessler, D. M. T. (2002). Windfall and socially distributed willpower: The psychocultural dynamics of rotating savings and credit associations in a Bengkulu village. Ethos, 30, 25-48.

Frank, R. H. (2000). Luxury fever: Money and happiness in an era of excess. Princeton, NJ: Princeton University Press.

Frank, R. H. (2004). What price the moral high ground? Ethical dilemmas in competitive environments. Princeton, NJ: Princeton University Press.

Gerber, A. S., Green, D. P., and Larimer, C. W. (2008). Social pressures and voter turnout: Evidence from a large-scale field experiment. American Political Science Review, 102, 33-48.

Gneezy, U., and Rustichini, A. (2000). A fine is a price. Journal of Legal Studies, 29, 1-17.

Goldstein, N. J., Griskevicius, V., and Cialdini, R. B. (2007). Invoking social norms: A
Harnessing the power of reputation

social psychology perspective on improving hotels’ linen-reuse programs. *Cornell Hospitality Quarterly, 48*, 145-150.

Griskevicius, V., Cantú, S. M., and Van Vugt, M. (2012). The evolutionary bases for sustainable behavior: Implications for marketing, policy, and social engineering. *Journal of Public Policy and Marketing, 31*, 1-14.

Griskevicius, V., Tybur, J.M., Gangestad, S.W., Perea, E. F., Shapiro, J. R., and Kenrick, D. T. (2009). Aggress to impress: Hostility as an evolved context-dependent strategy. *Journal of Personality and Social Psychology, 96*, 980-994.

Griskevicius, V., Tybur, J. M., Sundie, J. M., Cialdini, R. B., Miller, G. F., and Kenrick, D. T. (2007). Blatant benevolence and conspicuous consumption: When romantic motives elicit strategic costly signals. *Journal of Personality and Social Psychology, 93*, 85-102.

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*, 392-404.

Gurven, M., Allen-Arave, W., Hill, K., and Hurtado, A. M. (2000). “It’s a Wonderful Life”: Signaling generosity among the Ache of Paraguay. *Evolution and Human Behaviour, 21*, 263-282.

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

Hall, M., Lasby, D., Ayer, S., and Gibbons, W. D. (2009). *Caring Canadians, involved Canadians: Highlights from the 2007 Canada Survey of Giving, Volunteering and Participating* (Catalogue Number 71-542-X). Ottawa, ON: Statistics Canada.

Hamilton, W. D. (1964). The genetical evolution of social behaviour (I and II). *Journal of Theoretical Biology, 7*, 1-52.

Harbaugh, W. T. (1998). What do donations buy? A model of philanthropy based on prestige and warm glow. *Journal of Public Economics, 67*, 269-284.

Hardin, G. 1968 The tragedy of the commons. *Science, 162*, 1243-1248.

Hardy, C., and Van Vugt, M. (2006). Giving for glory in social dilemmas: The competitive altruism hypothesis. *Personality and Social Psychology Bulletin, 32*, 1402-1413.

Harris, A. C., and Madden, G. J. (2002). Delay discounting and performance on the prisoner’s dilemma game. *The Psychological Record, 52*, 429-440.

Hess, N. H., Hagen, E. H. (2006). Psychological adaptations for assessing gossip veracity. *Human Nature, 17*, 337-354.

Hoffman, E., McCabe, K., Schachat, K., and Smith, V. (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior, 7*, 346-380.

Horita, Y. (2010). Punishers may be chosen as providers but not as recipients. *Letters on Evolutionary Behavioral Science, 1*, 6-9.

Iredale, W., Van Vugt, M., and Dunbar, R. (2008). Showinf off in humans: Male generosity as a mating signal. *Evolutionary Psychology, 6*, 386-392.

Kaplan, H., and Hill, K. (1985). Food sharing among Ache foragers: Tests of explanatory hypotheses. *Current Anthropology, 26*, 223-246.

Karan, D., and McConnell, M. A. (2012). Hey look at me: The effect of giving circles on
Harnessing the power of reputation

giving (Working Paper No. 17737). The National Bureau of Economic Research, Cambridge, MA. Retrieved from http://www.nber.org/papers/w17737.

Kazdin, A. E. (2009). Psychological science’s contributions to a sustainable environment: Extending our reach to a grand challenge of society. American Psychologist, 64, 339-356.

Kelly, S., and Dunbar, R. I. M. (2001). Who dares, wins: Heroism versus altruism in women’s mate choice. Human Nature, 12, 89-105.

Keser, C. (2003). Experimental games for the design of reputation management systems. IBM Systems Journal, 42, 498-506.

Kimura, A., Mukawa, N., Yamamoto, M., Masuda, T., Yuasa, M., Goto, S. I., . . . Wada, Y. (2012). The influence of reputational concerns on purchase intention of fair-trade foods among young Japanese adults. Food Quality and Preference, 26, 204-210.

Kiyonari, T., and Barclay, P. (2008). Cooperation in social dilemmas: Free-riding may be thwarted by second-order rewards rather than punishment. Journal of Personality and Social Psychology, 95, 826-842.

Knafo, A., and Plomin, R. (2006). Parental discipline and affection and children’s prosocial behavior: Genetic and environmental links. Journals of Personality and Social Psychology, 90, 147-164.

Kokko, H., Johnstone, R. A., and Clutton-Brock, T. H. (2001). The evolution of cooperative breeding through group augmentation. Proceedings of the Royal Society of London B, 268, 187-196.

Kümmerli, R., Colliard, C., Fiechter, N., Petitpierre, B., Russier, F., and Keller, L. (2007). Human cooperation in social dilemmas: Comparing the Snowdrift game with the prisoner’s dilemma. Proceedings: Biological Sciences, 274, 2965-2970.

Kurzban, R., and Houser, D. (2005). Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations. PNAS, 102, 1803-1807.

Low, B. S., and Heinen, J. T. (1993). Population, resources, and environment: implications of human behavioral ecology for conservation. Population and Environment, 15, 7-41.

Masclet, D., Noussair, C., Tucker, S., and Villevial, M. C. (2003). Monetary and nonmonetary punishment in the voluntary contributions mechanism. The American Economic Review, 93, 366-380.

Mifune, N., Hashimoto, H., and Yamagishi, T. (2010). Altruism towards in-group members as a reputation mechanism. Evolution and Human Behavior, 31, 109-117.

Milinski, M., Semman, D., Bakker, T. C. M., and Krambeck, H. J. (2001). Cooperation through indirect reciprocity: image scoring or standing strategy? Proceedings of the Royal Society B, 268, 2495-2501.

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., Krambeck, H. J., and Marotzke, J. (2006). Stabilizing the Earth’s climate is not a losing game: Supporting evidence from public goods experiments. Proceedings of the National Academy of Sciences USA, 103, 394-3998.
Harnessing the power of reputation

Mulder, L. B., van Dijk, E., DeCremer, D., and Wilke, H. A. M. (2006). Undermining trust and cooperation: The paradox of sanctioning systems in social dilemmas. *Journal of Experimental Social Psychology, 42*, 147-162.

Nelissen, R. (2008). The price you pay: Cost-dependent reputation effects of altruistic punishment. *Evolution and Human Behavior, 29*, 242-248.

Nesse, R. M. (2005). Natural selection and the regulation of defenses: A signal detection analysis of the smoke detector principle. *Evolution and Human Behavior, 26*, 88-105.

Nowak, M. A., and Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature, 437*, 1291-1298.

Oda, R., Niwa, Y., Honma, A., and Hiraishi, K. (2011). An eye-like painting enhances the expectation of a good reputation. *Evolution and Human Behavior, 32*, 166-171.

Oliver, P. (1980). Rewards and punishments as selective incentives for collective action: Theoretical investigations. *American Journal of Sociology, 85*, 1356-1375.

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge, UK: Cambridge University Press.

Ostrom, E., Walker, J., and Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *American Political Science Review, 86*, 404-417.

Piazza, J., and Bering, J. M. (2008). Concerns about reputation via gossip promote generous allocations in an economic game. *Evolution and Human Behavior, 29*, 172-178.

Price, M. E. (2005). Punitive sentiment among the Shuar and in industrialized societies: Cross-cultural similarities. *Evolution and Human Behavior, 26*, 279-287.

Rege, M., and Telle, K. (2004). The impact of social approval and framing on cooperation in public good situations. *Journal of Public Economics, 88*, 1625-1644.

Rigdon, M., Ishii, K., Watabe, M., and Kitayama, S. (2009). Minimal social cues in the dictator game. *Journal of Economic Psychology, 30*, 358-367.

Roberts, G. (1998). Competitive altruism: From reciprocity to the handicap principle. *Proceedings of the Royal Society B, 265*, 427-431.

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

Rockenbach, B., and Milinski, M. (2006). The efficient interaction of indirect reciprocity and costly punishment. *Nature, 444*, 718-723.

Searcy, W. A., and Nowicki, S. (2005). *The evolution of animal communication: Reliability and deception in signaling systems*. Princeton, NJ: Princeton University Press.

Seinen, I., and Schram, A. (2006). Social status and group norms: Indirect reciprocity in a helping experiment. *European Economic Review, 50*, 581-602.

Semmann, D., Krambeck, H. J., and Milinski, M. (2004). Strategic investment in reputation. *Behavioral Ecology and Sociobiology, 56*, 248-252.

Sexton, S. E., and Sexton, A. L. (2011). Conspicuous conservation: The Prius effect and willingness to pay for environmental bona fides. Manuscript in preparation. Retrieved from [http://works.bepress.com/sexton/11](http://works.bepress.com/sexton/11).

Shang, J., and Croson, R. (2006). The impact of social comparisons on nonprofit fundraising. *Research in Experimental Economics, 11*, 143-156.

Simpson, B., and Willer, R. (2008). Altruism and indirect reciprocity: The interaction of
person and situation in prosocial behavior. Social Psychology Quarterly, 71, 37-52.
Singer, T., Seymour, B., O’Doherty, J. P., Stephan, K. E., Dolan, R. J., and Frith, C. D.
(2006). Empathic neural responses are modulated by the perceived fairness of
others. Nature, 439, 466-469.
Smith, E. A. (2004). Why do good hunters have higher reproductive success? Human
Nature, 15, 343-364.
Smith, E. A., and Bliege Bird, R. (2000). Turtle hunting and tombstone opening: Public
generosity as costly signaling. Evolution and Human Behavior, 21, 245-262.
Soetevent, A. R. (2005). Anonymity in giving in a natural context: A field experiment in 30
churches. Journal of Public Economics, 89, 2301-2323.
Sosis, R. (2004). The adaptive value of religious ritual. American Scientist, 92, 166-172.
Sparks, A., and Barclay, P. (submitted). Habituation to a false cue to observation: Eye
images increase generosity, but not for long. Manuscript submitted for publication.
Sylwester, K., and Roberts, G. (2010). Cooperators benefits through reputation-based
partner choice in economic games. Biology Letters, 6, 659-652.
Tinbergen, N. (1968). On war and peace in animals and man. Science, 160, 1411-1418.
Todd, P. M. (2001). Fast and frugal heuristics for environmentally bounded minds. In G.
Gigerenzer and R. Selten (Eds.), Bounded rationality: The adaptive toolbox (pp. 51-
70). Cambridge, MA: MIT Press.
Trivers, R. (1971). The evolution of reciprocal altruism. Quarterly Review of Biology, 46,
35-57.
Veblen, T. (1899/1994). The theory of the leisure class. Penguin Classics.
Wedekind, C., and Milinski, M. (2000) Cooperation through image scoring in humans.
Science, 288, 850-852.
West, S. A., El Mouden, C., and Gardner, A. (2011). Sixteen misconceptions about the
evolution of cooperation in humans. Evolution and Human Behaviour, 32, 231-262.
Wilson, M., and Daly, M. (2004). Do pretty women inspire men to discount the future?
Biology Letters, 271, S177-S179.
Xiao, E., and Houser, D. (2009). Avoiding the sharp tongue: Anticipated written message
promote fair economic exchange. Journal of Economic Psychology, 30, 393-404.
Yamagishi, T. (1986). The provision of a sanctioning system as a public good. Journal of
Personality and Social Psychology, 51, 110-116.
Zahavi, A., and Zahavi, A. (1997). The handicap principle: A missing piece of Darwin’s
puzzle. New York: Oxford University Press.