

1 Chapter 10

---

2 **The evolution of charitable behaviour**  
3 **and the power of reputation**

4 Pat Barclay

5 **Introduction**

6 Humans are arguably the most cooperative species on the planet when it comes to non-kin inter-  
7 actions. Humans regularly help non-kin in both formal and informal settings. For example, in my  
8 home country of Canada in 2007 alone, 84% of people made charitable donations for a value of  
9 almost \$10 billion, and 46% of people volunteered an average of 166 hours (Hall et al. 2009). Of  
10 course these figures are underestimates of the real total, as they do not include other ubiquitous  
11 forms of informal help such as favours, advice, exchange of benefits, restaurant tips, or restraint  
12 from overharvesting resources and selfish competition. Although helping rates certainly vary,  
13 Canada is far from alone in such behaviour.

14 From an evolutionary perspective, such charitable behaviour is very puzzling. Why would an  
15 organism ever do something to benefit another if it were costly to do so? Shouldn't such behav-  
16 iour be naturally selected against? If non-helpers save themselves the cost of helping, we might  
17 expect them to have an advantage in competition against those who possess more charitable sen-  
18 timents. Yet many organisms do help others, and the existence of such behaviour is one of the  
19 central puzzles in the study of evolution and behaviour (e.g. Vogel 2004; Pennisi 2009).

20 It has been several decades since scientists realized that behaviour does not evolve 'for the good  
21 of the species' (Hamilton 1964; Williams 1966; Dawkins 1976), and much theoretical and empir-  
22 ical work has identified the selective pressures that can select for and maintain such apparently  
23 selfless behaviour. Given that this is a chapter on *applied* evolutionary psychology, I will not  
24 attempt a formal review of all models of non-kin cooperation, but will instead present the basics  
25 on some of them before proceeding to potential applications and limitations on those applica-  
26 tions. Many of these applications are tentative, either because there has not yet been formal  
27 research on their use in applied charitable settings, or in the case of using reputational pressures,  
28 because the theories themselves are still incomplete and first require more basic research.  
29 Nevertheless, it is hoped that the following will give evolutionary researchers some ideas on how  
30 to apply their research and will give non-evolutionary applied researchers a reason to start think-  
31 ing about evolutionary psychology.

32 **Avoiding persistent confusion and misunderstandings**

33 In evolutionary theory, a psychology for helping will evolve whenever possessing such a psychol-  
34 ogy returns net benefits in terms of inclusive fitness. Several processes could cause that to happen,  
35 all of which ultimately rely on either direct benefits to self or indirect benefits to close relatives  
36 (Reeve 2000; Foster et al. 2006; West et al. 2007). Whatever form these benefits take, they function  
37 to reinforce the behaviour. This reinforcement can cause a person to learn to be generous,

1 for example by internalization of the norms which promote helping. Over evolutionary time,  
2 such reputational benefits can cause the evolution of cooperative sentiment. In the following sec-  
3 tions I will outline some of the major selective pressures that could have selected for helping, and  
4 will speculate on some possible applications of this knowledge for increasing the amount of  
5 charitable behaviour observed.

6 It is very easy to misunderstand evolutionary approaches and predictions, given that evolution-  
7 ary thinkers use different terminology and ask different types of questions than other fields. Even  
8 among evolutionists, many bitter arguments have resulted from nothing more than differences in  
9 terminology or levels of analysis. As such, it is well worthwhile to clarify some basic points before  
10 proceeding, as this can help avoid later misunderstandings.

### 11 **What phenomena am I referring to?**

12 By ‘charitable behaviour’, I refer to all types of behaviour that benefit others at a cost to self,  
13 regardless of whether the helper later receives benefits for doing so.<sup>1</sup> Different disciplines use dif-  
14 ferent terms for behaviours that benefit others at a cost to self, including ‘altruism’, ‘cooperation’,  
15 ‘generosity’, and ‘pro-social behaviour’. Because of these differing definitions, these terms  
16 are somewhat loaded with unintended implications and controversies. Rather than debate the  
17 ‘correct’ use of such terms, here I will attempt to use the more neutral ‘helping’ as a general pur-  
18 pose term for all behaviours that benefit others at a cost to self, and I will simply invite readers to  
19 substitute whatever term in their field is most appropriate for such behaviours.

20 I am lumping together many types of helping because evolutionary psychology usually  
21 treats different types of help with the same theoretical models and experimental games,  
22 whether that help is the reciprocal exchange of aid, coalitional support, emergency helping,  
23 courtship giving, tipping, charitable donations, conservation of resources, production of public  
24 goods, or pro-environmental behaviour (to name just a few). Different types of help occur in  
25 different currencies with different costs and benefits and will be performed by different  
26 people towards different recipients, but from an evolutionary perspective there is no a priori  
27 reason to predict that different types of help will respond in *qualitatively* different ways to changes  
28 in such parameters. As such, evolutionary psychology generally assumes some generality  
29 of its models.

### 30 **Levels of analysis**

31 Most people would agree that many body parts exist because they have an evolutionary function.  
32 For example, hearts pump blood, eyes extract visual information about the world, and the visual  
33 cortex processes that information to make it usable. The presence of these organs increases sur-  
34 vival, and hence, reproductive success. Behaviour can also affect survival and reproductive suc-  
35 cess, so it is also subject to natural selection, along with whatever psychological mechanisms and  
36 cognitive processes produce the behaviour. Behaviour is not inherited directly—it is produced by  
37 brains, which in turn are produced by developmental interactions between genes and environ-  
38 ments. If a gene interacts with the developmental environment in such a way as to grow a brain  
39 which produces beneficial behaviour more often, then that gene will be selected for because of  
40 those positive consequences.

41 If a particular behaviour reliably results in a specific positive outcome, then the production  
42 of that outcome can be said to be the ‘function’ of that behaviour. For example, the function of

---

<sup>1</sup> I am defining ‘costs’ as relative to the global population, not just relative to one’s local group (West et al. 2006). For a discussion of this important distinction, see West et al. (2007).

1 eating is the acquisition of energy and essential nutrients. Of course, most of us don't think about  
2 energy acquisition when we eat: we think about our hunger. However, the reason that hunger  
3 exists is because it leads us to eat, which results in the acquisition of energy and nutrients. Thus,  
4 there is a difference between the psychological mechanism (hunger) and the function of that  
5 mechanism (acquisition of energy): the evolutionary function is the reason that the mechanism  
6 exists. If we look at charitable behaviour, we must similarly distinguish between the psychological  
7 mechanisms (e.g. empathy, guilt, concern for reputation) and their potential evolutionary func-  
8 tions (e.g. acquisition of social benefits such as future reciprocation); we experience only the  
9 mechanism, but its function is the reason that the mechanisms exists.

10 In fact, there are four different types of question that we need to distinguish between in order  
11 to fully understand behaviour (Tinbergen 1963). We must ask about *psychological mechanisms*,  
12 and ask whether helping is caused by empathy, a warm glow, conformity to norms, or something  
13 else within the individual (i.e. 'What went on in my brain that made *me* help?'). Secondly, we  
14 must ask about the *development* of those mechanisms, and whether they are present at birth,  
15 learned via reinforcement or internalization of norms, or otherwise acquired within an individu-  
16 al's developmental history (i.e. 'How did my genes and environment interact to produce that  
17 psychological mechanism?'). Thirdly, we must ask about *evolutionary function*, and whether help-  
18 ing exists because it invites reciprocation, brings social prestige, wards off punishment, or raises  
19 inclusive fitness in other ways (i.e. 'What are the benefits of the behaviour, i.e. why does *anyone*  
20 have that psychological mechanism and developmental response?'). Finally, we must understand  
21 the *phylogeny* or *evolutionary history* of different helping behaviours, whether they are common  
22 to all our ancestors or are recently derived, and from what (i.e. 'How and when did that psycho-  
23 logical mechanism evolve?'). Most behavioural scientists seek to understand the psychological  
24 mechanisms and their development, but we must also understand why those mechanisms exist at  
25 all and how they evolved. Thus, these four types of questions are complementary, not mutually  
26 exclusive. Maternal helping behaviour is a good illustrative example. Maternal concern (*mecha-*  
27 *nism*) gets triggered in certain circumstances and involves specific neurobiological processes.  
28 Mothers are capable of such concern because of the way their genes and past social environments  
29 interacted (*development*). Maternal concern is present in all primates and many mammals, and  
30 uses similar brain pathways, indicating that it is evolutionarily ancient (*evolutionary history*). This  
31 concern exists because it causes mothers to help their offspring and thus contributes to offspring  
32 survival and reproduction (*function*). All four questions must be addressed for a complete under-  
33 standing.

34 It is very easy to confuse these different levels of analysis, especially when dealing with helping  
35 behaviour. No one would suggest that people are always consciously concerned with energy  
36 acquisition whenever they eat, but many have assumed that people are consciously concerned  
37 with the acquisition of benefits whenever they help (e.g. conscious concern for reputation) or  
38 they have interpreted other researchers as making that argument (e.g. Fehr and Henrich 2003;  
39 Gintis et al. 2003). This has led to much confusion in the literature. Instead, we must realize that  
40 mechanisms, development, function, and phylogeny are simply four different and complemen-  
41 tary questions. By investigating function and phylogeny instead of just mechanism and develop-  
42 ment, it gives us a greater understanding of helping behaviour and how to promote it, for example  
43 by changing our social circumstances to ensure that the function is fulfilled.

#### 44 **Adaptive on average**

45 Evolutionary researchers propose that helping behaviour exists because it increases inclusive fit-  
46 ness on average. However, not every instance of helping will do so, nor does it need to. For a trait

1 to evolve, it needs to bring its bearer more net benefits on average than a different trait would  
2 (or would have done so in ancestral environments), even if that trait occasionally results in instances  
3 of non-beneficial behaviours. For example, empathy might have evolved because it causes people  
4 to behave better towards associates, which in turn results in the empathic person receiving better  
5 behaviour (Frank 1988). However, it might also cause people to sometimes help total strangers  
6 who will never be in a position to reciprocate. If the overall benefits from helping the 'right' people  
7 (e.g. friends and associates) outweigh the cost of helping both 'right' and 'wrong' people (e.g. the  
8 occasional stranger), then being empathic pays better than being non-empathic, so empathy will  
9 evolve. Of course, rather than using a fixed 'help' or 'no-help' strategy, evolution refines our psy-  
10 chological mechanisms to make adaptive distinctions between different people and situations, for  
11 example by making us more likely to help kin versus non-kin (Daly and Wilson 1988; Stewart-  
12 Williams 2007), friends versus strangers (e.g. Majolo et al. 2006), attractive people versus unat-  
13 tractive people (Farrelly et al. 2007), and non-competitors versus competitors (West et al. 2006).

14 Despite this fine-tuning of our psychological adaptations, there is no evolutionary theory that  
15 predicts that every single instance of helping should increase fitness, because errors are inevitable  
16 in any decision-making process (Haselton and Buss 2000; Nesse 2005). Rather than focus on  
17 specific instances of helping and asking whether they are adaptive, one must focus on the *mecha-*  
18 *nism*: is this particular 'decision rule' adaptive? Does it bring more benefits in the real world than  
19 an alternative decision rule? By increasing your likelihood of helping, you simultaneously increase  
20 the likelihood of helping in the 'right' circumstances (i.e. when you can benefit from it) and of  
21 helping in the 'wrong' circumstances (i.e. when you cannot benefit); the optimal level of helping  
22 strikes a balance between the former and the latter (see Figure 10.1). If the 'right' circumstances  
23 are much more frequent, or if failing to help in those circumstances is very costly (e.g. lost oppor-  
24 tunities for reciprocity or other reputation), then our psychological mechanisms will cause us to  
25 err on the side of more helping. This will occasionally result in helping when it brings no benefits,  
26 but this cost is outweighed by the overall benefit. Conversely, as the 'wrong' circumstances  
27 increase in frequency, or as the cost of helping in those circumstances increases, then we should  
28 err on the side of low helping. This will result in more missed opportunities to receive benefits,  
29 but these are outweighed by the overall lower costs. Thus, although errors are inevitable, an evo-  
30 lutionary approach helps to predict whether people will err on the side of helping too much  
31 versus too little, and this knowledge can be used to increase people's likelihood of helping.

## 32 **Inclusive fitness and vested interests**

### 33 **Inclusive fitness theory**

34 Any gene can propagate copies of itself either by aiding the reproduction of its host or the repro-  
35 duction of those who are statistically likely to carry copies of that same gene, especially close kin  
36 (Hamilton 1964). This is known as inclusive fitness theory, and it predicts that all else being equal,  
37 people will be more likely to help those who are close kin, as opposed to distant kin or non-  
38 relatives. This prediction has been abundantly confirmed in non-humans and in humans from  
39 many societies (for one textbook review, see Barrett et al. 2002). Inclusive fitness theory is the  
40 cornerstone of modern studies of evolution and behaviour. All else being equal, people can expect  
41 much better treatment from kin than non-kin, and interactions with kin are generally more coop-  
42 erative than interactions with non-kin in the same situations (Daly and Wilson, 1988). Despite its  
43 importance, I will spend little time discussing kinship because: 1) it is much more obvious to  
44 most evolutionists, and 2) helping non-kin is more difficult to understand and some aspects  
45 remain unresolved.

		Decision	
		Help	Don't help
<p>Helping would increase fitness (e.g. recipient is kin/friend, people notice, vested interest in outcome, etc.)</p> <p>Helping wouldn't increase fitness (e.g. recipient is stranger, no one sees, etc.)</p>	<p><b>Cost:</b> pay to help</p> <p><b>Benefit:</b> reciprocity, gain reputation, direct benefit to self or to kin, etc.</p> <p><b>Total:</b> net benefit</p>	<p><b>Cost:</b> punishment, lose reputation, etc.</p> <p><b>Benefit:</b> none (though also no cost paid to help)</p> <p><b>Total:</b> net cost</p>	
	<p><b>Cost:</b> pay to help</p> <p><b>Benefit:</b> none</p> <p><b>Total:</b> net cost</p>	<p><b>Cost:</b> none</p> <p><b>Benefit:</b> none (though also no cost paid to help)</p> <p><b>Total:</b> no cost or benefit</p>	

**Fig. 10.1** Fitness costs and benefits for helping and not-helping. Given that all organisms cannot know with 100% certainty what circumstances they are in, errors are inevitable (i.e. occasionally helping in the 'wrong' circumstances or not-helping in the 'right' circumstances). Rather than look at the payoffs for specific outcomes, we need to look at the overall payoff for helping versus not-helping averaged across all 'right' and 'wrong' circumstances (weighted by their frequency). Our psychological mechanisms and emotions evolved to be adaptive on average: they will make us err on the side of helping when the *average* payoff for helping outweighs the *average* payoff for not-helping (and vice versa), which in turn depends on the magnitude of the costs/benefits for each outcome and the frequency of the 'right' circumstances.

1 Before leaving the topic, it is worth mentioning that kinship need not be real to have an effect  
 2 on helping: presenting people with unconscious cues of kinship can be enough to trigger coop-  
 3 erative behaviour. For example, facial resemblance is a cue of kinship (DeBruine 2005), and  
 4 DeBruine (2002) showed that laboratory participants entrusted more money to people whose  
 5 pictures had been 'morphed' to slightly resemble their own faces. Using a similar technique,  
 6 Krupp and colleagues (2008) showed that people became increasingly more cooperative as the  
 7 number of perceived kin in the group increased. Oates and Wilson (2002) found that people were  
 8 more willing to help people who had the same family name as themselves, especially if it was a rare  
 9 name (and thus a more reliable cue of kinship). Salmon (1998) found that the use of kinship  
 10 terms in political speeches was effective at getting people to agree with the speaker, especially for  
 11 firstborns who are typically closer to their families. This research all suggests that cues of kinship  
 12 can be effective at eliciting charitable behaviour, whether the cues are verbal, facial, nominal, or  
 13 merely kinship terms. Of course, overt use of such fictive kinship cues might trigger hostile reac-  
 14 tions from those who are aware of their function, just as my evolutionary classmates in graduate  
 15 school and I used to get angry at one particular campus political persona who would refer to us as  
 16 'brothers and sisters'.

### 17 Vested interests

18 Kinship is not the only way for one individual to have a fitness stake in the well-being of another  
 19 individual (Roberts 2005). For example, people have a vested interest in the well-being of their

1 friends and allies, and in the continued existence of their social groups (Tooby and Cosmides  
2 1996; Kokko et al. 2001; Lahti and Weinstein 2005). If these friends or social groups were to die,  
3 disappear or disband, then those involved would be worse off. As such, it is sometimes worth  
4 incurring an unreciprocated cost in order to ensure their well-being. Even extreme generosity  
5 towards one's in-group can be adaptive if the alternative is extinction at the hands of hostile out-  
6 groups (Reeve and Hölldobler 2007; Barclay and Benard, submitted). Many cases of apparent  
7 'altruism'—such as so-called 'altruistic punishment' of non-cooperators in a group—are in fact  
8 beneficial to the person performing the act (West et al. 2007).

9 Most perspectives would predict that people will display more charitable behaviour if they have  
10 a vested interest in the outcome, such as cancer victims donating towards cancer research.  
11 Evolutionary theory reminds us that people can also have a vested interest in the well-being of  
12 other people or even entire groups, and that unreciprocated helping should occur towards such  
13 recipients. Political leaders are well aware of this when helping allies or bailing out failing compa-  
14 nies that perform necessary national services. Anything that increases the interdependency of  
15 people can also increase people's stake in one another's well-being, and thus the charitable behav-  
16 iour observed towards others.

## 17 **Direct reciprocity**

18 Trivers (1971) noted that two individuals who exchange help can be better off than individuals  
19 who neither give nor receive help. His paper inspired much empirical and theoretical work on the  
20 evolution of pairwise reciprocal exchange of aid, which was most famously demonstrated by the  
21 success of the 'Tit-for-Tat' cooperative strategy in Axelrod's computer simulations on the evolu-  
22 tion of cooperation (Axelrod 1984). The general conclusions from this broad literature are that it  
23 pays to help others if there is a chance that they can reciprocate, that many people are 'conditional  
24 cooperators' who are willing to cooperate if and only if others follow suit, and that cooperation is  
25 more likely to arise when the costs are low, the benefits are high, and the chance of interacting  
26 again is high.

## 27 **Structural solutions to enhance direct reciprocity**

28 Since the costs, benefits, and probability of future interaction all affect the evolution of coopera-  
29 tion, changing these simple factors can have strong effects on the rates of helping. Changing  
30 actual costs and benefits may be difficult, but sometimes all that is necessary is to correct misper-  
31 ceptions about these factors. The stability of partnerships and probability of future interaction  
32 (or perceptions thereof) can also sometimes be changed: by making individuals more likely to  
33 interact again, it makes it more worthwhile for them to engage in reciprocal exchange of aid.

34 A less-known but very important structural solution is to change the scale of competition  
35 between people (West et al. 2006). West and colleagues used a cooperative game (a Prisoner's  
36 Dilemma) where people were divided into subgroups of three within a class, and they could coop-  
37 erate within those subgroups to earn points. They gave a prize to either the highest-scorer in each  
38 subgroup or to the highest scorers within the whole class. When the prize went to the highest-  
39 scorers in the class, there was an incentive for people to cooperate within the subgroups in order  
40 to better compete against the rest of the class ('global competition'), so cooperation flourished.  
41 However, when the prize went to the highest-scorer within each subgroup, there was less reason  
42 to cooperate within that subgroup because doing so would directly aid one's competitors ('local  
43 competition'), and cooperation plummeted. This has many implications for the structure  
44 of human groups (Crespi 2006): many business and academic groups are set up such that people  
45 compete for promotions or grades against the same people they could be cooperating with.

1 Such competitive structures will inhibit cooperation. If instead groups were designed to shift  
2 competition from a local scale (within-group) to a global scale (against the broader population),  
3 then it could greatly increase the level of within-group cooperation. One way to do this is  
4 to eliminate incentive programmes that rate people only relative to their local group (e.g. choos-  
5 ing or promoting only the best member of each group), because this local competition can  
6 encourage hostility and sabotaging within the groups instead of cooperation (see also Sober and  
7 Wilson 1998).

## 8 **Conditional cooperation and expectations of others' behaviour**

9 Some of the most obvious ways to promote charitable behaviour involve people's evolved tenden-  
10 cies towards reciprocation and conditional cooperation. This can involve inducing reciprocation,  
11 creating perceptions that others are cooperating, and reducing the negative effects of others who  
12 fail to help. For example, many charities induce reciprocity by giving small gifts, thus inducing a  
13 feeling of obligation to reciprocate. Other charities announce a large capital contribution  
14 (Andreoni 2006), which is like having someone else make the first cooperative move in a coop-  
15 erative venture.

16 Evolutionary researchers have noted that people can occasionally make errors by failing to help  
17 when they intend to, or by misinterpreting whether others have helped. These errors reduce help-  
18 ing because they trigger retaliation from partners, who assume that the non-cooperation occurred  
19 and was deliberate (e.g. Axelrod 1984; Nowak and Sigmund 1992). As such, successful coopera-  
20 tive strategies must be able to recover from these errors by being forgiving of the occasional defec-  
21 tion (e.g. Nowak and Sigmund 1992) or by being slightly more generous than one's partner so  
22 that one's errors don't make one appear stingy (e.g. Van Lange et al. 2002; Van der Bergh and  
23 Dewitte 2006). As such, error-correction mechanisms can help maintain cooperation, as will  
24 promoting forgiveness of the occasional failure to help.

25 Evolutionary theory predicts that organisms will try to avoid cooperating with those who do  
26 not cooperate (Trivers 1971; Axelrod 1984; Cosmides and Tooby 1992), so any changes in the  
27 perceived or expected cooperativeness of others should change people's likelihood of coopera-  
28 tion. There is much evidence—much of it not explicitly evolutionary—to support this prediction  
29 (for a review, see Pruitt and Kimmel 1977). For example, people are much more likely to cooper-  
30 ate if given the opportunity to communicate with each other beforehand (e.g. Caldwell 1976;  
31 Ostrom et al. 1992; Davis and Holt 1993); this can occur either because they can make promises  
32 to each other or simply because communication gives them an opportunity to assess others'  
33 cooperativeness. People are more likely to cooperate with members of an in-group or when in-  
34 group membership is made salient (e.g. Messick and Brewer 1983), and Toshio Yamagishi and  
35 colleagues have shown that this is because people have higher expectations of cooperation from  
36 in-group members (Yamagishi and Kiyonari 2000; Yamagishi 2003). Furthermore, the threat of  
37 punishment increases people's contributions towards a group fund because of both a direct effect  
38 on people who can receive the punishment and an indirect effect on the resulting expectations of  
39 cooperation from others (Shinada and Yamagishi 2007).

40 There are multiple ways to change people's expectations about others' cooperation. Changing  
41 the structural parameters (costs, benefits, length of interactions) should do so. People are more  
42 cooperative if given the opportunity to gradually build and escalate trust rather than immediately  
43 commit to high helping (e.g. Roberts and Sherratt 1998; Kurzban et al. 2001; Roberts and Renwick  
44 2003). People are also expected to be sensitive to the perceived frequency of helpers and non-  
45 helpers in the population (e.g. McNamara and Houston 2002; Barclay 2008; McNamara et al.  
46 2009), so one could work to change this perception or correct any misperceptions. Observing  
47 non-cooperation has a greater effect on people's helping behaviour than observing cooperation

1 (Monteresso et al. 2002; Offerman 2002), so it is better to focus attention on the cooperators  
2 rather than the non-cooperators.

3 One relatively simple way to change people's expectations about others' cooperation is to  
4 change how situations are presented to them. If different framings can change what is considered  
5 appropriate, then they will change not only what people feel they 'should' do, but will also change  
6 people's expectations about others' behaviour and thus make it 'safe' to cooperate. For example,  
7 people are more likely to donate money to their group if the group is framed as a 'community'  
8 and the money is framed as belonging primarily to the community (Rege and Telle 2004).  
9 Similarly, people are more likely to cooperate with partners when a cooperative scenario is framed  
10 as the 'Community Game' versus the 'Wall Street Game' (Lieberman et al. 2004). The practical  
11 implications of this are straightforward: pay attention to presentation. Further research should  
12 investigate the relative effects of different types of pro-social framings, for example whether pre-  
13 senting helping as a 'social obligation' is more effective than presenting it as a 'nice thing to do',  
14 and whether very obvious framings will be perceived as being manipulative and cause people to  
15 react negatively.

## 16 **Generalized helping and public reputation**

17 Humans are especially known for their willingness to help people outside of their immediate  
18 circle of kin, friends, and allies. This psychology is often believed to have evolved because of  
19 various forms of reputational costs and benefits, such as indirect reciprocity, punishment, and  
20 costly signalling; I have previously reviewed these in greater detail elsewhere (Barclay, in press).  
21 There is considerable overlap of predictions and applications arising from these, so I will  
22 discuss applications from them together in the following section after briefly introducing each in  
23 this section.

### 24 **Indirect reciprocity**

25 Direct reciprocity occurs when individuals exchange helping acts, but sometimes such acts  
26 are reciprocated by someone other than the recipient—this is termed 'indirect reciprocity'  
27 (Alexander 1987) and is reviewed in detail by Nowak and Sigmund (2005). In the best-known  
28 form of indirect reciprocity (termed 'downstream reciprocity'), those who help tend to acquire a  
29 good reputation, and this increases the probability that someone else will help them when they  
30 need it. If someone is observed refusing to help, this harms their reputation and the chances that  
31 others will help them. Such indirect reciprocity can evolve provided that reputations can be  
32 tracked, and people distinguish between a justified failure to help a 'bad' person and an unjusti-  
33 fied failure to help a 'good' person (see Nowak and Sigmund 2005). Much experimental evidence  
34 shows that people who give help are more likely to receive help, even from those who have never  
35 received anything from them and never will (e.g. Gurven et al. 2000; Wedekind and Milinski  
36 2000; Milinski et al. 2001; Seinen and Schram 2006).

37 In addition to this 'downstream reciprocity', it is worth briefly mentioning one other form of  
38 indirect reciprocity called 'upstream reciprocity', where those who are helped become more likely  
39 to help others whom they later encounter (Nowak and Sigmund 2005). This is popularly known  
40 as 'paying it forward', and can evolve because it triggers direct reciprocity from those helped  
41 (Nowak and Roch 2007). Experimental evidence shows that after people receive help, they are  
42 indeed more likely to help others that they later encounter (Bartlett and DeSteno 2006; Fowler  
43 and Christakis 2010); similar effects have also been found in rats (Rutte and Taborsky 2007).  
44 Although this 'upstream reciprocity' is not actually related to reputation—people are basing their  
45 helping decisions on their own receipt of help rather than the reputation of others—it still deserves



1 mention because it too can be harnessed to promote charitable behaviour. If people are more  
2 likely to help others after receiving help, then one simple idea is to start a wave of charitable  
3 behaviour and simply let it propagate itself through social networks (Christakis and Fowler 2009).  
4 If the structure of a social network can be changed to better propagate these waves, so much the  
5 better.

## 6 **Punishment**

7 Indirect reciprocity shows that the rewards of a good reputation can sustain charitable behaviour.  
8 Having a bad reputation can have explicit costs too, as people are very willing to pay to impose  
9 costs upon those who do not pitch in their fair share in cooperative situations (e.g. Yamagishi  
10 1986; Ostrom et al. 1992). People will seemingly punish even if they themselves receive no benefit  
11 from the punishee's future cooperative behaviour (Fehr and Gächter 2002). The presence of pun-  
12 ishment makes it costly to refuse to help, thus providing a selective pressure for helping (e.g. Boyd  
13 and Richerson 1992). Whereas rewards and indirect reciprocity are efficient for eliciting charita-  
14 ble behaviour from just a few group members, punishment is more efficient if cooperation is  
15 required from all members of a population, because one only needs to use sanctions on rare non-  
16 cooperators (Oliver 1980). Fostering such peer-to-peer punishment may be one way to promote  
17 unanimous cooperation.

18 Despite its power to promote cooperation, peer-to-peer punishment is theoretically puzzling  
19 because all group members benefit when an uncooperative group member is coerced into coop-  
20 erating, but only the punisher pays the cost to deliver such peer-to-peer punishment (Oliver  
21 1980; Yamagishi 1986). As such, the evolutionary function of such apparently spiteful behaviour  
22 is currently being hotly debated. For example, it could be maintained by punishers receiving a  
23 good or a tough reputation (Brandt et al. 2003; Barclay 2006), conformist imitation of group  
24 members (e.g. Guzman et al. 2007), or by punishers or their kin directly benefiting from the sub-  
25 sequent increased group cooperation (e.g. West et al. 2007). An understanding of the relative  
26 importance of these forces in maintaining peer-to-peer punishment will increase people's ability  
27 to regulate the cooperation in their own groups.

## 28 **Costly signalling**

### 29 **Informational value**

30 Helping others is costly, especially for those who do not have the resources to spare. In addition,  
31 those with little concern for others (and no intent to engage in later cooperation) will be particu-  
32 larly unwilling to incur the costs of helping because they receive no benefits from doing so. Thus,  
33 those who observe someone else helping can infer certain traits about the helper, such as his/her  
34 abilities, resources, and/or cooperative intent (Hawkes 1991; Boone 1998; Smith and Bliege Bird  
35 2000; Gintis et al. 2001; Lotem et al. 2002). Helping thus functions as an honest signal of those  
36 traits, and the costs of helping make it not worth the cost for those who do not honestly possess  
37 those traits. Observers benefit from pairing, mating with, or befriending those who have high  
38 abilities, resources, and cooperative intent. As such, it is beneficial to be seen to help because this  
39 makes one a more desirable social partner or mate (e.g. Smith et al. 2003; Barclay 2004, 2010a).  
40 This explanation of helping is but one example of 'costly signalling', an idea which has been of  
41 great use in evolutionary biology for explaining exaggerated traits such as peacocks' tails and  
42 extravagant displays (for reviews, see Zahavi and Zahavi 1997; Searcy and Nowicki 2005). Applying  
43 costly signalling theory to explain human helping does not require that people are aware of what  
44 their generosity signals (though obviously some are), only that the reputational benefits can sus-  
45 tain the behaviour.



### 1 Evidence for costly signalling

2 Field and laboratory studies both provide evidence that helping functions as a signal of personal  
3 traits and that others respond to it as such. Hunting big game and widely sharing the meat has  
4 been described as a costly signal of physical abilities, and data show that good hunters have higher  
5 reproductive success and more extra-marital affairs than poorer hunters—even in societies where  
6 the hunters receive no more meat than anyone else (e.g. Kaplan and Hill 1985; Hawkes 1991;  
7 Smith et al. 2003; Smith 2004). Laboratory studies show that heroic male risk-takers are more  
8 desirable mates than non-heroes (Kelly and Dunbar 2001; Farthing 2005); similar effects have  
9 been found with other types of helping (Barclay 2010a). Hosting large feasts is an effective way to  
10 advertise one's resources because viewers have a vested interest in attending to the signal, not only  
11 because of its informational value but also to receive some of the food and gifts being distributed  
12 (e.g. Boone 1998; Smith and Bliege Bird 2000). Large monetary donations have also been inter-  
13 preted as a means to acquire prestige (Harbaugh 1998).

14 With the former types of help, these acts signal the ability or resources to provide help that  
15 not everyone can provide. Helping can also signal concern for others, and can thus be a cue of  
16 one's likelihood of cooperating or helping in the future. Less theoretical work has been done on  
17 helping as an honest signal of cooperative intent and how the honesty of such signalling is  
18 maintained, but few people will be surprised to learn that helping at time A predicts later helping  
19 at time B (e.g. Clark 2002; Kurzban and Houser 2005); in fact, this is the basis of personality.  
20 Research also suggests that people treat helping as though it carries informational value.  
21 For example, people are more trusting of others who give money to other group members  
22 (Wedekind and Braithwaite 2002; Barclay 2004, 2006), to charity (Albert et al. 2007), or to their  
23 previous social partners (Keser 2003). Costly signalling theory has only recently been used to  
24 explain helping behaviour, so it is not yet clear when helping should signal qualities that not eve-  
25 ryone possesses (e.g. physical abilities or monetary resources) versus behavioural traits such as a  
26 willingness to cooperate; though see André (2010) and Barclay and Reeve (submitted) for recent  
27 thoughts on this.

28 Some researchers argue that religious rituals may serve as costly signals towards members of  
29 one's religious group (see Sosis 2004). The cost of such rituals deters those who would take the  
30 benefits of group membership without actually valuing the group's position. This idea is impor-  
31 tant because religious donations comprise the largest type of all charitable donations (Hall et al.  
32 2009). Hazing (initiation rituals involving humiliation, harassment, or abuse) may serve a similar  
33 function in fraternities, another group known for high rates of solidarity and monetary support  
34 from members.

### 35 When do we expect costly signalling?

36 When should observers pay most attention to signals of quality or cooperative intent? This is an  
37 important question, because it affects the size of the reputational benefits, and thus the effective-  
38 ness of reputation at increasing helping behaviour. Audiences should only attend to signals if they  
39 are informative, i.e. they carry information that is useful and is honest on average. If everyone has  
40 the same abilities or resources, then audiences gain no useful information from attending to  
41 costly signals of these qualities (Gintis et al. 2001). Attending to such signals only pays if there are  
42 enough low quality individuals to make them worth avoiding. Similarly, if everyone is coopera-  
43 tive, then there is no reason to demand costly signals of cooperative intent from potential partners  
44 because it is easier to simply accept all potential partners; it only pays to discriminate if there is a  
45 risk of being cheated on (McNamara and Houston 2002; McNamara et al. 2009). Thus, as varia-  
46 tion in a population increases in terms of abilities, resources, and/or cooperative intent, more  
47 people will attend to helping behaviour as a cue of those traits, in turn increasing the reputational



1 benefits for helping. Thus, reputation will be more effective at increasing charitable behaviour  
2 when there is high variation in the population.

### 3 **The power of reputation**

#### 4 **Preconditions and evidence**

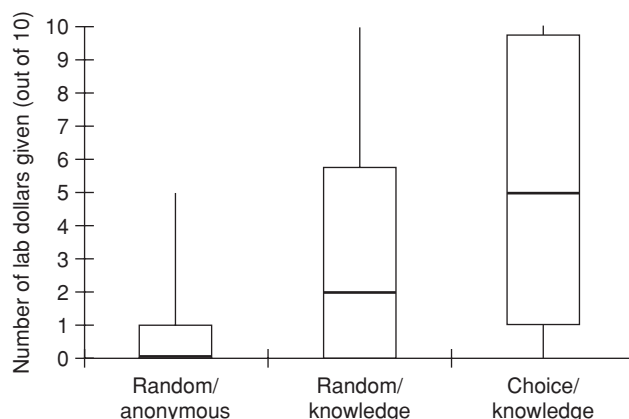
5 For cooperative sentiment to evolve via indirect reciprocity, punishment, or costly signalling,  
6 people must be able to acquire a reputation for their helping (or conversely, not helping),  
7 and these reputational benefits must outweigh the costs of providing help. These benefits can  
8 come from the recipients of the help, those who observed it, or from others who hear about it.  
9 Indirect reciprocity theory and costly signalling theory both make many similar predictions, to  
10 the point that disentangling them is often difficult (and possibly unnecessary: see André 2010).  
11 However, both of these theories—as well as theories of punishment—make the prediction  
12 that helping behaviour should correlate with the magnitude and certainty of the reputational  
13 consequences.

14 These three theories predict that people will be more willing to help others when they can  
15 acquire a reputation for doing so, and this prediction has been abundantly confirmed. For exam-  
16 ple, people give more money to others in experimental games when their behaviour is publicly  
17 announced (Rege and Telle 2004; Bereczkei et al. 2007), made available to people they will later  
18 interact with (Milinski et al. 2002; Barclay 2004; Hardy and Van Vugt 2006), talked about or com-  
19 mented on by fellow participants (Masclét et al. 2003; Piazza and Bering 2008), or even just  
20 known to the experimenters (Hoffman et al. 1994). People are more likely to contribute to a fund  
21 to educate others about climate change if they can acquire a reputation for this giving, and they  
22 tend to be rewarded for doing so (Milinski et al. 2006). In an experimental tax game, participants  
23 were less likely to cheat on their taxes if the pictures of tax-evaders were to be made public, and  
24 this effect was in addition to any deterrent effects of monetary sanctions (Coricelli et al. 2010).  
25 Sometimes this behaviour is quite strategic, in that the same people help when they can acquire a  
26 reputation but then shortly afterwards refuse to help in anonymous situations (Milinski et al.  
27 2002; Semmann et al. 2004; Barclay and Willer 2007; see also Figure 10.2). However, adaptive  
28 responses to cues of reputation need not be conscious, and in real life are probably often driven  
29 by heightened emotions like guilt and shame in public.

#### 30 **Competitive altruism**

31 When people can choose with whom to interact, people will even actively compete to give more  
32 than others, in what has been termed ‘competitive altruism’ (Roberts 1998; Barclay 2004; Hardy  
33 and Van Vugt 2006; Barclay and Willer 2007; Sylwester and Roberts, 2010). Such behaviour  
34 increases one’s access to desirable social partners. Helping behaviour is sensitive to social com-  
35 parison, which makes sense in the light of this market-based competition over cooperative part-  
36 ners. For example, people give more money to radio fundraising if given information other  
37 people’s high donors (Shang and Croson, 2006), and are influenced by the perceived number of  
38 other people who have made pro-environmental decisions like conserving water by re-using hotel  
39 towels (for review, see Kazdin 2009).

40 Competitive altruism explains such why such social comparisons are effective: no one wants to  
41 lose out on partnership opportunities for being slightly less cooperative than others. Competitive  
42 altruism also explains how cooperative norms can escalate over time, as people try to outdo each  
43 other in terms of relative generosity. Fostering competitive giving could increase charitable  
44 behaviour, such as by naming university buildings after the highest donors. Other ways to



**Fig. 10.2** Experimental evidence for competitive giving. Participants in an experimental game gave little money to their partners when no one finds out how much they gave, gave more money when observed by a third party who might then interact with them, and gave the most when that third-party observer could choose with whom to interact (thus giving an incentive to appear more cooperative than one's partner). Redrawn from Barclay and Willer (2007).

1 foster competitive giving involve increasing people's ability to choose their social partners,  
 2 thus allowing the best cooperators to pair with the best cooperators and giving everyone an incen-  
 3 tive to be more helpful in order to pair up with better partners. Private clubs may use this by  
 4 allowing opportunities for high donors to meet and interact with each other; political fundraising  
 5 events (e.g. '\$500 a plate' dinners) can serve a similar function.

## 6 Implicit cues of reputation

7 People are so sensitive about reputation that cues of reputation do not even need to be explicit. In  
 8 one famous study, Haley and Fessler (2005) gave people the option of giving money to others in  
 9 an experimental game, and then varied whether there were a set of stylized eyespots on the com-  
 10 puter screen versus the laboratory's logo. Even though participants' decisions were completely  
 11 anonymous in either case, they gave more money when eyes were present than when absent.  
 12 These results have been replicated in other studies with slightly different experimental games and  
 13 stimuli (Burnham and Hare 2007; Mifune et al. 2010; though see the discussion in the later sec-  
 14 tion 'Situational potency of reputational cues'). Rigdon and colleagues (2009) were even able to  
 15 show this effect with three dots presented in a 'watching-eyes configuration' (downward triangle)  
 16 versus a neutral configuration (upward triangle). As evidence that this effect works outside the  
 17 laboratory, Bateson and colleagues (2006) found that people gave more money to pay for coffee  
 18 on an 'honour system' when there were eyes above the money jar than when there were flowers.  
 19 These studies all show that people's reputational concerns can be triggered to increase charitable  
 20 behaviour, and the triggers can be relatively minor.

21 People can also be influenced by being primed about other aspects of their reputation. For  
 22 example, after being put into a mating mindset by viewing pictures of attractive people and imag-  
 23 ining a perfect date with one of them, men and women reported being more willing to engage in  
 24 helping acts that could signal their generosity, prestige, or heroism (Griskevicius et al. 2007). Being  
 25 primed to think about their social status increases people's reported willingness to buy eco-friendly  
 26 products (e.g. hybrid cars) relative to similarly priced luxury products (Griskevicius et al. 2010).

1 If such effects can occur outside the laboratory, then many types of helping behaviour could be  
2 affected by triggering the right reputational concerns.

### 3 **Harnessing the power of reputation**

4 If we can harness these powers of reputation, then we can potentially create powerful and  
5 long-lasting increases in charitable behaviour. Evolutionary models of cooperation show  
6 that cooperation is more likely when people are likely to interact again or when a person's reputa-  
7 tion is more likely to come back to affect them. Currently, many interactions in developed coun-  
8 tries are with strangers, such that there is less likelihood of one's reputation benefiting or harming  
9 oneself in the long run. On the one hand, this is a drawback, but on the other hand it is a  
10 great opportunity: if we can somehow increase the likelihood that people's reputation will follow  
11 them to situations where it matters, then we can greatly increase cooperation. This could  
12 be accomplished by creating stable groups that are likely to have repeated interactions or by  
13 increasing the transmission of reputational information. The Internet may be particularly useful  
14 for this information transfer, especially the increased use of Internet social networking sites.  
15 Although we value privacy, some things like energy use affect others, so perhaps we should con-  
16 sider making such information more public so that reputational pressures will entice people to act  
17 more sustainably.

18 We can also use people's reputational concerns to foster demand for 'green' or socially respon-  
19 sible products. Once an act (such as purchasing socially responsible products) becomes associated  
20 as being 'good', systems of indirect reciprocity will tend to support that act and reward those who  
21 do it (Milinski et al. 2002; Panchanathan and Boyd 2004). As long as companies follow their eco-  
22 nomic interests by responding to consumer demands, then this can influence what products get  
23 produced, and can potentially reduce unsustainable or exploitative production. To phrase it  
24 poetically: 'money talks, but reputation helps determine what that money says' (Barclay 2010b).  
25 Consumer boycotts have been successful before, but an outright boycott is not necessary because  
26 in the long run, the larger market share will go to whichever companies or products are being  
27 bought more. Rather than wait for long-term effects, companies should be made aware of why  
28 people are avoiding their products, because this will result in much faster changes in their prac-  
29 tices. It is debatable whether it is more effective to focus on the reputations of individual consum-  
30 ers versus the reputations of Chief Executive Officers (CEOs) or other corporate decision-makers  
31 (Jennifer Jacquet, pers. comm. 19 June 19, 2010): focusing on the former can change the market  
32 forces such that it is in a company's interest to become socially responsible, whereas focusing on  
33 the latter has a greater per capita immediate effect because CEOs have more influence than the  
34 average consumer.

### 35 **Creating reputational pressures**

36 It is one thing to affect the anonymity of people's acts, but is another to make sure that reputa-  
37 tional forces will actually act to promote helpful or socially responsible behaviour (i.e. instead of  
38 people simply ignoring a given behaviour). Beyond making acts more public, how do we ensure  
39 that the right reputational pressures will be there to promote socially responsible acts? This is  
40 much easier asked than answered. Once an act is perceived as 'good', then those who actually are  
41 'good' will tend to do it, thus sustaining the correlation between that act and good character, such  
42 that reputation can promote the act. But what creates this perception in the first place? Unless  
43 that perception already exists, we need a way to link certain acts (e.g. buying only socially respon-  
44 sible products) with a 'good' reputation. Doing so will involve changing public perceptions to  
45 make some acts or purchases seem socially desirable and other acts seem socially undesirable,

1 but this is not always easily accomplished and there has not been much evolutionary research on  
2 this question.

3 Although it may be sufficient to simply assert that ‘good people do X and bad people do Y’, it is  
4 probably more effective to point out exactly how X benefits others and Y hurts others. Common  
5 knowledge of these benefits and harms is probably important, because it allows people to infer  
6 something about the character of whoever is performing the act. Someone who is genuinely con-  
7 cerned for others’ welfare will not readily do something that harms another. As such, anyone who  
8 knowingly performs such a harmful act can be inferred to not have such concern—but only if  
9 they should reasonably know that the act is socially harmful. One potential example of how the  
10 norms have changed is that of campaigns against drinking and driving: these have been very  
11 successful at changing social norms, such that in many areas the social pressures are more effec-  
12 tive at reducing impaired driving than police enforcement. By focusing on the social dangers of  
13 impaired driving, it changed people’s perceptions of those who drink and drive and made it  
14 socially unacceptable to do so. If it becomes more commonly known that certain acts (e.g. pollu-  
15 tion, unsustainable consumption) are socially harmful, then people will be more willing to sanc-  
16 tion those who do them. In Canada, anti-smoking advertisements are beginning to focus on  
17 smoking’s harm to others rather than to oneself, which subtly changes the social pressures. In a  
18 more extreme case, one can imagine a sort of ‘Asshole Campaign’ to shame those whose behav-  
19 iour negatively affects others, such as the use of ‘Gas guzzling = Asshole’ bumper stickers on  
20 highly inefficient vehicles.

21 One can attempt to link harmful behaviours with negative reputation by focusing on the social  
22 costs, or one can link positive behaviours with good reputation by focusing on the social benefits.  
23 Thus, education about positive effects can foster a reputation system to support that behaviour.  
24 For example, someone who cares about others’ welfare and has the opportunity to help at low cost  
25 will not knowingly refrain from doing. As such, whether or not someone performs a helping act  
26 will become a cue about that person’s concern for others—as long as the helper and the observer  
27 both know that the act actually does help others. One example of this is how Milinski and col-  
28 leagues (2006) showed that educating people about the dangers of climate change helps increase  
29 the social pressure to fight against climate change. In that experiment, education without reputa-  
30 tion had marginal effects on people’s willingness to fight climate change, but the combination of  
31 education and reputation was particularly effective. Thus, educating people about the effects of  
32 their behaviour is not separate from reputational effects. Instead, they are linked: education is a  
33 crucial component in creating the reputational pressures to sustain charitable behaviour, and  
34 reputational forces give people a reason to care about what they have learned and to act on it.

### 35 **Limitations**

36 Despite the potential for reputation to increase charitable behaviour, it has several important  
37 limitations that need to be overcome in order to promote sustained helping. If we attempt to use  
38 reputations to increase charitable behaviour without first giving thought to these issues, the net  
39 effect could be marginal, transitory, or possibly even counterproductive. For each of the following  
40 limitations—which are presented in increasing order of importance—I also attempt to provide  
41 solutions.

#### 42 **Costs of helping may outweigh reputational benefits**

43 It is easy to make the simplistic argument that charitable behaviour will increase when people can  
44 acquire a reputation. However, that prediction is based on people receiving social benefits for  
45 helping others, so we must actually assess whether the expected reputational benefits will out-  
46 weigh the costs of helping. When the expected reputational benefits are small relative to the

1 expected costs (e.g. when most people do not associate a particular helping act as being ‘good’),  
2 then reputation should have a negligible effect on helping, and may even decrease helping (see  
3 later section ‘Explicit incentives reduce intrinsic motivations for helping’). This point seems  
4 extraordinarily obvious, but is also surprisingly easy to forget once we try to incorporate multiple  
5 kinds of delayed costs and benefits, multiple observers, and competition over reputational bene-  
6 fits.

### 7 **Situational potency of reputational cues**

8 Sometimes reputation manifests itself with strong and tangible consequences, such as monetary  
9 sanctions (e.g. Yamagishi 1986) or the presence of mating opportunities (Griskevicius et al. 2007;  
10 Barclay 2010a). However, other cues—such as the use of eyespots to subconsciously trigger repu-  
11 tational concerns (e.g. Haley and Fessler 2005)—are much more subtle and are often not actually  
12 followed by real opportunities for reputation. The presence of reputation and non-anonymity  
13 may have effects across situations, but weak cues of reputation will be more easily overwhelmed  
14 by other features of the situation such as internalized norms regarding appropriate behaviour or  
15 other costs and benefits. Consistent with this idea, two recent studies have found that weak repu-  
16 tational cues did little to change giving behaviour in economic games which had strong situational  
17 cues and norms: offers and responses in Ultimatum Games are unaffected by the presence of oth-  
18 ers in a room (Lamba and Mace 2010), and amounts reciprocated in a Trust Game are unaffected  
19 by eyespots on a computer (Fehr and Schneider 2010). In these cases, the normative expectations  
20 of fairness—and internalization of these norms—may have outweighed the relatively weak repu-  
21 tational cues in those experiments. Thus, reputation may be more effective at eliciting coopera-  
22 tion in some situations than others, such that stronger reputational incentives are necessary in  
23 situations where there are other countervailing factors.

### 24 **Consequences of non-reinforcement**

25 People may be predisposed to be sensitive to their reputation, but they are also excellent learners  
26 whose behaviour is shaped by reinforcement and non-reinforcement. Many animals eventually  
27 habituate or acclimatize to stimuli that carry no useful information, and they cease performing  
28 non-reinforced behaviours (e.g. Domjan and Burkhard 1993). As such, we should expect humans  
29 to similarly habituate to stimuli that lose their usefulness in predicting the presence of important  
30 reputational opportunities, and responding to uninformative cues should be extinguished. For  
31 example, the presence of eyespots are a subtle cue of being watched and thus increase cooperation  
32 (e.g. Haley and Fessler 2005; Burnham and Hare 2007), but if eyespots are repeatedly presented  
33 to people without any tangible reputational opportunities, then people should eventually come to  
34 ignore them—just as predators eventually come to ignore the fake eyespots flashed by prey spe-  
35 cies to scare off predators (Stevens 2005). Consistent with this prediction, Soetevent (2005) found  
36 that reputational cues had only temporary effects on charitable donations in churches. To resolve  
37 this problem, would-be social engineers should ensure that when they cue reputational concerns,  
38 those cues should be occasionally followed by real opportunities for non-trivial reputational ben-  
39 efits. Otherwise, the cues will eventually lose their effectiveness with repeated use.

### 40 **Not everyone cares about reputation**

41 Some experiments show a non-zero proportion of people who continue to act selfishly even when  
42 they are being conspicuously observed. For example, Rege and Telle (2004) found that people  
43 donate more money towards a group project when they have to announce their contributions to  
44 the entire group than when such donations are anonymous, but the proportion of people giving  
45 zero is unchanged by the presence of non-anonymity. It is as if there is some low proportion of

1 'rugged individualists' or psychopaths who care little for the respect or disdain of others. If some  
2 people do not need or desire a reputation for helping, then it is not worth the cost for them to help  
3 others (Barclay and Reeve submitted). For example, in attractiveness terms, some people are  
4 already so desirable that they can get away with being selfish (Takahashi et al. 2006).

5 These persistent cheaters may drag down the cooperation of others, given that most people  
6 cooperate only if others also do so (Fischbacher et al. 2001). Furthermore, some types of coop-  
7 eration require unanimity (Oliver 1980), and would thus fail to work if there were some people  
8 who cared little about social pressures and rewards. When unanimous cooperation is required,  
9 tangible punishment of non-cooperators may be required to ensure that everyone cooperates  
10 (Oliver 1980), but that punishment must be strong enough to be effective (e.g. Gneezy and  
11 Rustichini 2000a).

### 12 Devaluing the future benefits of reputation

13 Investing in future reputational benefits is an investment in the future, so it only pays off indi-  
14 viduals who will be around to reap the long-term benefits. Humans and other organisms value  
15 the future less than the present, such that there is a strong temptation to avoid the immediate  
16 costs of helping others because future benefits are perceived as less valuable (Frank 1988; Stephens  
17 2000; Stephens et al. 2002). Supporting this idea, people who place more value on the future are  
18 more likely to cooperate with each other in experimental games (Harris and Madden 2002).

19 This devaluing of future benefits is potentially problematic. However, the degree to which peo-  
20 ple value the future is not static and can vary adaptively according to local circumstances and life  
21 history (Wilson and Daly 1997, 2004). Thus, increasing people's 'time horizons' is predicted to  
22 increase their charitable behaviour. This will take more than rhetorical exhortations for people to  
23 invest in the future; it will take the presence of cues indicating that investing in the future will  
24 actually pay off and that each given person is likely to be around for it. Such cues can include cues  
25 of long life expectancy, better futures, or realistic life options (Wilson and Daly 1997), though  
26 obviously these are major structural changes that are more easily proposed than implemented.

### 27 Explicit incentives reduce intrinsic motivations for helping

28 Whereas some people only help others when there are explicit incentives for doing so, others have  
29 internalized norms about helping and seem to have an intrinsic desire to help others (Simpson  
30 and Willer 2008). Some have argued that it is this intrinsic motivation which causes much human  
31 cooperation (e.g. Fehr and Fischbacher 2004). Unfortunately, much research shows that the pres-  
32 ence of extrinsic rewards can undermine intrinsic motivation for tasks (for reviews, see Deci et al.  
33 1999; Bowles 2008). Similarly, when punishment of non-cooperators is present, it undermines  
34 trust and reciprocity, possibly because it becomes unclear whether others are cooperating out of  
35 intrinsic motivation or out of fear of punishment (Fehr and Rockenbach 2003; Mulder et al.  
36 2006). This reduction in intrinsic motivation can even outweigh any increased cooperation due  
37 to explicit incentives: two classic studies found that the presence of fines made parents *more* likely  
38 to be late to pick up their children from daycare than when no fines were present (Gneezy and  
39 Rustichini 2000a) and that giving small rewards for good performance on IQ tests makes people  
40 perform *worse* than if they receive no rewards at all (Gneezy and Rustichini 2000b). Bowles (2008)  
41 reviews a number of similar cases.

42 This undermining of intrinsic motivation is a very serious problem that must be considered  
43 whenever using incentives for helping. One potential solution is to rely on implicit incentives,  
44 such as verbal rewards, which do not undermine intrinsic motivation (Deci et al. 1999). Even bet-  
45 ter would be to simply create reputational opportunities for people to apply their own informal  
46 social sanctions: these can still be effective at changing behaviour (Barr 2001; Fessler, 2002;



1 Masclot et al. 2003), but they do not rely on the type of top-down explicit institutional incentives  
2 which would undermine intrinsic motivation to cooperate. Eventually, these informal social pres-  
3 sures can cause people to internalize the desire to help. Explicit or institutional incentives could  
4 be reserved for situations where reputational pressures are insufficient (see earlier section ‘Not  
5 everyone cares about reputation’).

#### 6 Identifying the reputational benefits may reduce them

7 If helping is designed to signal one’s good character, then it may be detrimental to call attention  
8 to any benefits that people receive. Doing so makes it unclear whether people are helping because  
9 they are concerned for others or because they are merely concerned with their own reputation—  
10 people tend to like the former type of person and be ambivalent or hostile towards the latter. Some  
11 people may even use this ambiguity to downplay the generosity of others: by calling attention to  
12 the reputational benefits someone else receives for helping, it limits the reputational advantage he/  
13 she receives. This is an example of ‘do-gooder derogation’ (Monin 2007) that is expected when  
14 competition over reputation is high, and is especially expected from non-cooperators who seek to  
15 justify their lack of helping (and thus prevent a reputational loss).

16 This problem is presently tricky to resolve because there is not much theoretical or empirical  
17 research on signals of cooperative intent, and even less research on how the effectiveness of such  
18 signals varies as a function of the apparent costs and benefits (e.g. audience size). One solution to  
19 this problem might be to create subtle opportunities for people to benefit from a reputation, but to  
20 strictly avoid any mention of people’s reputational benefits. In this case, the social rewards and  
21 sanctions can occur on their own without it being commonly known that people do indeed benefit  
22 from helping others. A second solution would be to announce other people’s generosity on their  
23 behalf so they don’t have to personally flaunt it; this allows them to get a reputation without engag-  
24 ing in shameless self-promotion, and gives them an excuse for why their generosity becomes com-  
25 mon knowledge. Thirdly, even if people are explicitly helping to gain a good reputation, it is worth  
26 comparing such people to others who refuse to help regardless of their reputation. It is also worth  
27 comparing people who seek status via helping versus those who seek status via costly consumption:  
28 someone who conspicuously gives extra meat away is nicer than someone who conspicuously burns  
29 it (Gintis et al. 2001), someone with an expensive pro-environmental product is nicer than some-  
30 one with an equally expensive unsustainable product (Griskevicius et al. 2010), and someone who  
31 spends money to save the last whale is morally superior to someone who spends money to catch and  
32 eat that whale (Barclay 2010b). By using the appropriate comparisons, we can help ensure that the  
33 identification of reputational benefits does not completely eliminate those benefits.

## 34 Conclusion

35 An evolutionary approach focuses on the *function* of charitable behaviour. In addition to examin-  
36 ing the psychological mechanisms underlying helping and the development of those mechanisms,  
37 evolutionists ask why humans possess psychological mechanisms in the first place, and why  
38 humans develop those mechanisms instead of just learning to be selfish. In this chapter, I have  
39 outlined some of the major selective pressures responsible for the evolution of helping, and I have  
40 suggested means by which this knowledge can be used to promote charitable behaviour. Focusing  
41 on the ultimate costs and benefits of helping can lead to research to create situations in which  
42 charitable behaviour pays off in the long run. This does not necessarily mean that people con-  
43 sciously seek to benefit from helping others, but instead means that these benefits will increase the  
44 frequency of helping either via learning (including internalization of norms) or via natural selec-  
45 tion for cooperative sentiment.

1 In particular, I have focused on harnessing the power of reputation because it is a novel  
2 approach with great potential to increase helping behaviour towards non-kin. At the same time,  
3 I have attempted to point out some limitations that need to be overcome in order for reputational  
4 pressures to be effective. Without consideration of these limitations, harnessing the power of  
5 reputation can be ineffective, and at worst could even reduce helping—recall the daycare study by  
6 Gneezy and Rustichini, (2000a) discussed in the section ‘Explicit incentives reduce intrinsic moti-  
7 vations for helping’. Theories of general reputation are far from complete, and there are many  
8 issues that still require resolving, such as how the magnitude of reputational benefits are affected  
9 by the apparent costs of helping, the frequency of helpers, the timescale, the type of reward, and  
10 especially from the public knowledge that there *are* benefits to having a good reputation. Hopefully  
11 the identification of these issues will inspire future research and lead to a more mature science of  
12 reputation. And hopefully such a science of reputation will have great power to increase charita-  
13 ble behaviour in the world.

## 14 References

- 15 Albert, M., Güth, W., Kirchler, E., and Maciejovsky, B. (2007). Are we nice(r) to nice(r) people? An  
16 experimental analysis. *Experimental Economics*, **10**, 53–69.
- 17 Alexander, R.D. (1987). *The biology of moral systems*. Aldine de Gruyter, New York.
- 18 André, J.B. (2010). The evolution of reciprocity: social types or social incentives? *The American Naturalist*,  
19 **175**, 197–210.
- 20 Andreoni, J. (2006). Leadership giving in charitable fund-raising. *Journal of Public Economic Theory*,  
21 **8**, 1–22.
- 22 Axelrod, R. (1984). *The evolution of cooperation*. Basic Books, New York.
- 23 Barclay, P. (2004). Trustworthiness and competitive altruism can also solve the ‘tragedy of the commons’.  
24 *Evolution and Human Behavior*, **25**, 209–20.
- 25 Barclay, P. (2006). Reputational benefits for altruistic punishment. *Evolution and Human Behavior*, **27**,  
26 325–44.
- 27 Barclay, P. (2008). Enhanced recognition of defectors depends on their rarity. *Cognition*, **107**, 817–28.
- 28 Barclay, P. (2010a). Altruism as a courtship display: some effects of third-party generosity on audience  
29 perceptions. *British Journal of Psychology*, **101**, 123–35.
- 30 Barclay, P. (2010b). ‘Harnessing the power of reputation.’ Invited talk at the TEDxGuelphU conference  
31 (independently licensed TED conference), April 2010, University of Guelph, Guelph, ON.
- 32 Barclay, P. (in press). *Reputation and the evolution of generosity*. Nova Science Publishers, Hauppauge, NY.
- 33 Barclay, P. and Benard, S. (submitted). ‘Power corrupts but competition for power corrupts more’:  
34 manipulation of perceived threats to preserve dominance in cooperative groups.
- 35 Barclay, P. and Reeve, H.K. (submitted). Extravagant and everyday generosity: the varying relationship  
36 between generosity and individual quality.
- 37 Barclay, P. and Willer, R. (2007). Partner choices creates competitive altruism in humans. *Proceedings of the*  
38 *Royal Society B*, **274**, 749–53.
- 39 Barr, A. (2001). *Social dilemmas and shame-based sanctions: experimental results from rural Zimbabwe*.  
40 University of Oxford, Institute of Economics and Statistics, Centre for the Study of African Economies,  
41 Oxford.
- 42 Barrett, L., Dunbar, R.I.M., and Lycett, J. (2002). *Human evolutionary psychology*. Princeton University  
43 Press, Princeton, NJ.
- 44 Bartlett, M.Y. and DeSteno, D. (2006). Gratitude and prosocial behavior: helping when it costs you.  
45 *Psychological Science*, **17**, 319–25.
- 46 Bateson, M., Nettle, D., and Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world  
47 setting. *Biology Letters*, **2**, 412–14.

AQ: Please update  
if available.



- 1 Bereczkei, T., Birkas, B., and Kerekes, Z. (2007). Public charity offer as a proximate factor of evolved  
2 reputation-building strategy: an experimental analysis of a real-life situation. *Evolution and Human*  
3 *Behavior*, **28**, 277–84.
- 4 Boone, J. L. (1998). The evolution of magnanimity: when is it better to give than to receive? *Evolution and*  
5 *Human Behavior*, **9**, 1–21.
- 6 Bowles, S. (2008). Policies designed for self-interested citizens may undermine ‘the moral sentiments’:  
7 evidence from economic experiments. *Science*, **320**, 1605–9.
- 8 Boyd, R. and Richerson, P.J. (1992). Punishment allows the evolution of cooperation (or anything else) in  
9 sizable groups. *Ethology and Sociobiology*, **13**, 171–95.
- 10 Brandt, H., Hauert, C., and Sigmund, K. (2003). Punishment and reputation in spatial public goods games.  
11 *Proceedings of the Royal Society B*, **270**, 1099–104.
- 12 Burnham, T. and Hare, B. (2007). Engineering human cooperation: does involuntary neural activation  
13 increase public goods contributions? *Human Nature*, **18**, 88–108.
- 14 Caldwell, M.D. (1976). Communication and sex effects in a five-person Prisoner’s Dilemma game. *Journal*  
15 *of Personality and Social Psychology*, **33**, 273–80.
- 16 Christakis, N.A. and Fowler, J.H. (2009). *Connected: the surprising power of our social networks and how they*  
17 *shape our lives*. Little, Brown and Company, New York.
- 18 Clark, J. (2002). Recognizing large donations to public goods: an experimental test. *Managerial and*  
19 *Decision Economics*, **23**, 33–44.
- 20 Coricelli, G., Joffily, M., Montmarquette, C., and Villeval, M.C. (2010). Cheating, emotions, and  
21 rationality: an experiment on tax evasion. *Experimental Economics*, **13**, 226–47.
- 22 Cosmides, L. and Tooby, J. (1992). Cognitive adaptations for social exchange. In: J. Barkow, L. Cosmides,  
23 and J. Tooby (eds), *The adapted mind: evolutionary psychology and the generation of culture*,  
24 pp. 163–228. Oxford University Press, New York.
- 25 Crespi, B. (2006). Cooperation: close friends and common enemies. *Current Biology*, **16**, R414–15.
- 26 Daly, M. and Wilson, M. (1988). *Homicide*. Aldine de Gruyter, New York.
- 27 Davis, D.D. and Holt, C.A. (1993). *Experimental economics*. Princeton University Press, Princeton, NJ.
- 28 Dawkins, R. (1976). *The selfish gene*. Oxford University Press, Oxford.
- 29 DeBruine, L.M. (2002). Facial resemblance enhances trust. *Proceedings of the Royal Society B*,  
30 **269**, 1307–12.
- 31 DeBruine, L.M. (2005). Trustworthy but not lust-worthy: context-specific effects of facial resemblance.  
32 *Proceedings of the Royal Society B*, **272**, 919–22.
- 33 Deci, E.L., Koestner, R., and Ryan, R.M. (1999). A meta-analytic review of experiments examining the  
34 effects of extrinsic rewards on internal motivation. *Psychological Bulletin*, **125**, 627–8.
- 35 Domjan, M., and Burkhard, B. (1993). *The principles of learning and behavior (3rd Edition)*. Brooks/Cole  
36 Publishing, Pacific Grove, CA.
- 37 Farrelly, D., Lazarus, J., and Roberts, G. (2007). Altruists attract. *Evolutionary Psychology*, **5**, 313–29.
- 38 Farthing, G.W. (2005). Attitudes toward heroic and nonheroic physical risk takers as mates and friends.  
39 *Evolution and Human Behavior*, **26**, 171–85.
- 40 Fehr, E. and Fischbacher, U. (2004). Social norms and human cooperation. *Trends in Cognitive Sciences*, **8**,  
41 185–90.
- 42 Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature*, **415**, 137–40.
- 43 Fehr, E. and Henrich, J. (2003). Is strong reciprocity a maladaptation? On the evolutionary foundations of  
44 human altruism. In: P. Hammerstein (ed.), *Genetic and cultural evolution of cooperation*, pp. 55–82.  
45 MIT Press, Cambridge, MA.
- 46 Fehr, E. and Rockenbach, B. (2003). Detrimental effects of sanctions on human altruism. *Nature*,  
47 **422**, 137–40.
- 48 Fehr, E. and Schneider, F. (2010). Eyes are on us, but nobody cares: are eye cues relevant for strong  
49 reciprocity? *Proceedings of the Royal Society B*, **277**, 1315–23.





- 1 Fessler, D.M.T. (2002). Windfall and socially distributed willpower: the psychocultural dynamics of  
2 rotating savings and credit associations in a Bengkulu village. *Ethos*, **30**, 25–48.
- 3 Fischbacher, U., Gächter, S., and Fehr, E. (2001). Are people conditionally cooperative? Evidence from a  
4 public goods experiment. *Economics Letters*, **71**, 397–404.
- 5 Foster, K.R., Wenseleers, T., Ratnieks, F.L.W., and Queller, D.C. (2006). There is nothing wrong with  
6 inclusive fitness. *Trends in Ecology and Evolution*, **21**, 599–600.
- 7 Fowler, J.H. and Christakis, N.A. (2010). Cooperative behavior cascades in human social networks.  
8 *Proceedings of the National Academy of Sciences of the USA*, **107**, 5334–8.
- 9 Frank, R.H. (1988). *Passions within reason*. Norton, New York.
- 10 Gintis, H., Bowles, S., Boyd, R., and Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution  
11 and Human Behavior*, **24**, 153–72.
- 12 Gintis, H., Smith, E.A., and Bowles, S. (2001). Cooperation and costly signaling. *Journal of Theoretical  
13 Biology*, **213**, 103–19.
- 14 Gneezy, U. and Rustichini, A. (2000a). A fine is a price. *Journal of Legal Studies*, **29**, 1–17.
- 15 Gneezy, U. and Rustichini, A. (2000b). Pay enough or don't pay at all. *The Quarterly Journal of Economics*,  
16 **3**, 791–810.
- 17 Griskevicius, V., Tybur, J.M., Sundie, J.M., Cialdini, R.B., Miller, G.F., and Kenrick, D.T. (2007). Blatant  
18 benevolence and conspicuous consumption: when romantic motives elicit strategic costly signals.  
19 *Journal of Personality and Social Psychology*, **93**, 85–102.
- 20 Griskevicius, V., Tybur, J.M., and Van den Bergh, B. (2010). Going green to be seen: status, reputation, and  
21 conspicuous conservation. *Journal of Personality and Social Psychology*, **98**, 392–404.
- 22 Gurven, M., Allen-Arave, W., Hill, K., and Hurtado, A.M. (2000). 'It's a Wonderful Life': signaling  
23 generosity among the Ache of Paraguay. *Evolution and Human Behaviour*, **21**, 263–82.
- 24 Guzman, R.A., Rodríguez-Sickert, C., and Rowthorn, R. (2007). When in Rome, do as the Romans do: the  
25 coevolution of altruistic punishment, conformist learning, and cooperation. *Evolution and Human  
26 Behavior*, **28**, 112–17.
- 27 Haley, K.J. and Fessler, D.M.T. (2005). Nobody's watching? Subtle cues enhance generosity in an  
28 anonymous economic game. *Evolution and Human Behavior*, **26**, 245–56.
- 29 Hall, M., Lasby, D., Ayer, S., and Gibbons, W.D. (2009). *Caring Canadians, involved Canadians:  
30 highlights from the 2007 Canada survey of giving, volunteering and participating*. Statistics Canada,  
31 Ottawa, ON.
- 32 Hamilton, W.D. (1964). The genetical evolution of social behaviour (I and II). *Journal of Theoretical  
33 Biology*, **7**, 1–52.
- 34 Harbaugh, W.T. (1998). What do donations buy? A model of philanthropy based on prestige and warm  
35 glow. *Journal of Public Economics*, **67**, 269–84.
- 36 Hardy, C. and Van Vugt, M. (2006). Giving for glory in social dilemmas: the competitive altruism  
37 hypothesis. *Personality and Social Psychology Bulletin*, **32**, 1402–13.
- 38 Harris, A.C. and Madden, G.J. (2002). Delay discounting and performance on the Prisoner's Dilemma  
39 game. *The Psychological Record*, **52**, 429–40.
- 40 Haselton, M.G. and Buss, D.M. (2000). Error management theory: a new perspective on biases in cross-sex  
41 mind reading. *Journal of Personality and Social Psychology*, **78**, 81–91.
- 42 Hawkes, K. (1991). Showing off: tests of an hypothesis about men's foraging goals. *Ethology and  
43 Sociobiology*, **12**, 29–54.
- 44 Hoffman, E., McCabe, K., Schachat, K., and Smith, V. (1994). Preferences, property rights, and anonymity  
45 in bargaining games. *Games and Economic Behavior*, **7**, 346–80.
- 46 Kaplan, H. and Hill, K. (1985). Food sharing among Ache foragers: tests of explanatory hypotheses. *Current  
47 Anthropology*, **26**, 223–46.
- 48 Kazdin, A.E. (2009). Psychological science's contributions to a sustainable environment: extending our  
49 reach to a grand challenge of society. *The American Psychologist*, **64**, 339–56.





- 1 Kelly, S. and Dunbar, R.I.M. (2001). Who dares, wins: heroism versus altruism in women's mate choice.  
2 *Human Nature*, **12**, 89–105.
- 3 Keser, C. (2003). Experimental games for the design of reputation management systems. *IBM Systems*  
4 *Journal*, **42**, 498–506.
- 5 Kokko, H., Johnstone, R.A., and Clutton-Brock, T.H. (2001). The evolution of cooperative breeding  
6 through group augmentation. *Proceedings of the Royal Society B*, **268**, 187–96.
- 7 Krupp, D.B., DeBruine, L.M., and Barclay, P. (2008). A cue of kinship promotes cooperation for the public  
8 good. *Evolution and Human Behavior*, **29**, 49–55.
- 9 Kurzban, R. and Houser, D. (2005). Experiments investigating cooperative types in humans: a complement  
10 to evolutionary theory and simulations. *Proceedings of the National Academy of Sciences of the USA*,  
11 **102**, 1803–7.
- 12 Kurzban, R., McCabe, K., Smith, V.L., and Wilson, B.J. (2001). Incremental commitment and reciprocity in  
13 a real-time public goods game. *Personality and Social Psychology Bulletin*, **27**, 1662–73.
- 14 Lahti, D.C. and Weinstein, B.S. (2005). The better angels of our nature: group stability and the evolution of  
15 moral tension. *Evolution and Human Behavior*, **26**, 47–63.
- 16 Lamba, S. and Mace, R. (2010). People recognise when they are really anonymous in an economic game.  
17 *Evolution and Human Behavior*, **31**, 271–8.
- 18 Liberman, V., Samuels, S.M., and Ross, L. (2004). The name of the game: predictive power of reputations  
19 versus situational labels in determining Prisoner's Dilemma game moves. *Personality and Social*  
20 *Psychology Bulletin*, **30**, 1175–85.
- 21 Lotem, A., Fishman, M.A., and Stone, L. (2002). From reciprocity to unconditional altruism through  
22 signaling benefits. *Proceedings of the Royal Society B*, **270**, 199–205.
- 23 Majolo, B., Ames, K., Brumpton, K., Garratt, R., Hall, K., and Wilson, N. (2006). Human friendship  
24 favours cooperation in the Iterated Prisoner's Dilemma. *Behaviour*, **143**, 1383–95.
- 25 Masclet, D., Noussair, C., Tucker, S., and Villeval, M.C. (2003). Monetary and nonmonetary punishment in  
26 the voluntary contributions mechanism. *The American Economic Review*, **93**, 366–80.
- 27 McNamara, J.M. and Houston, A.I. (2002). Credible threats and promises. *Philosophical Transactions of the*  
28 *Royal Society B*, **357**, 1607–16.
- 29 McNamara, J.M., Stephens, P.A., Dall, S.R.X., and Houston, A.I. (2009). Evolution of trust and  
30 trustworthiness: social awareness favours personality differences. *Proceedings of the Royal Society B*,  
31 **276**, 605–13.
- 32 Messick, D.M. and Brewer, M.B. (1983). Solving social dilemmas. In: L. Wheeler and P. Shaver (eds),  
33 *Review of personality and social psychology*, Vol. **4**, pp. 11–44. Sage Publishing, Beverly Hills, CA.
- 34 Mifune, N., Hashimoto, H., and Yamagishi, T. (2010). Altruism towards in-group members as a reputation  
35 mechanism. *Evolution and Human Behavior*, **31**, 109–17.
- 36 Milinski, M., Semmann, D., Bakker, T.C.M., and Krambeck, H.J. (2001). Cooperation through indirect  
37 reciprocity: image scoring or standing strategy? *Proceedings of the Royal Society B*, **268**, 2495–501.
- 38 Milinski, M., Semmann, D., and Krambeck, H.J. (2002). Reputation helps solve the 'tragedy of the  
39 commons'. *Nature*, **415**, 424–6.
- 40 Milinski, M., Semmann, D., Krambeck, H.J., and Marotzke, J. (2006). Stabilizing the Earth's climate is not a  
41 losing game: supporting evidence from public goods experiments. *Proceedings of the National Academy*  
42 *of Sciences of the USA*, **103**, 394–8.
- 43 Monin, B. (2007). Holier than me? Threatening comparison in the moral domain. *Revue Internationale de*  
44 *Psychologie Sociale*, **20**, 53–68.
- 45 Monterosso, J., Ainslie, G., Toppi Mullen, P. and Gault, B. (2002). The fragility of cooperation: a false  
46 feedback study of a sequential iterated prisoner's dilemma. *Journal of Economic Psychology*, **23**, 437–48.
- 47 Mulder, L.B., van Dijk, E., DeCremer, D., and Wilke, H.A.M. (2006). Undermining trust and cooperation:  
48 the paradox of sanctioning systems in social dilemmas. *Journal of Experimental Social Psychology*, **42**,  
49 147–62.





- 1 Nesse, R.M. (2005). Natural selection and the regulation of defenses: a signal detection analysis of the  
2 smoke detector principle. *Evolution and Human Behavior*, **26**, 88–105.
- 3 Nowak, M.A. and Roch, S. (2007). Upstream reciprocity and the evolution of gratitude. *Proceedings of the*  
4 *Royal Society B*, **274**, 605–9.
- 5 Nowak, M.A. and Sigmund, K. (1992). Tit for tat in heterogenous populations. *Nature*, **355**, 250–3.
- 6 Nowak, M.A. and Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, **437**, 1291–8.
- 7 Oates, K. and Wilson, M. (2002). Nominal kinship cues facilitate human altruism. *Proceedings of the Royal*  
8 *Society B*, **269**, 105–9.
- 9 Offerman, T. (2002). Hurting hurts more than helping helps. *European Economic Review*, **46**, 1423–37.
- 10 Oliver, P. (1980). Rewards and punishments as selective incentives for collective action: theoretical  
11 investigations. *American Journal of Sociology*, **85**, 1356–75.
- 12 Ostrom, E., Walker, J., and Gardner, R. (1992). Covenants with and without a sword: self-governance is  
13 possible. *American Political Science Review*, **86**, 404–17.
- 14 Panchanathan, K. and Boyd, R. (2004). Indirect reciprocity can stabilize cooperation without the second-  
15 order free rider problem. *Nature*, **432**, 499–502.
- 16 Pennisi, E. (2009). On the origin of cooperation. *Science*, **325**, 1196–9.
- 17 Piazza, J. and Bering, J.M. (2008). Concerns about reputation via gossip promote generous allocations in an  
18 economic game. *Evolution and Human Behavior*, **29**, 172–8.
- 19 Pruitt, D.G. and Kimmel, M.J. (1977). Twenty years of experimental gaming: critique, synthesis, and  
20 suggestions for the future. *Annual Reviews of Psychology*, **28**, 363–92.
- 21 Reeve, H.K. (2000). Multi-level selection and human cooperation. *Evolution and Human Behavior*,  
22 **21**, 65–72.
- 23 Reeve, H.K. and Hölldobler, B. (2007). The emergence of a superorganism through  
24 intergroup competition. *Proceedings of the National Academy of Sciences of the USA*, **104**, 9736–40.
- 25 Rege, M. and Telle, K. (2004). The impact of social approval and framing on cooperation in public good  
26 situations. *Journal of Public Economics*, **88**, 1625–44.
- 27 Rigdon, M., Ishii, K., Watabe, M., and Kitayama, S. (2009). Minimal social cues in the dictator game.  
28 *Journal of Economic Psychology*, **30**, 358–67.
- 29 Roberts, G. (1998). Competitive altruism: from reciprocity to the handicap principle. *Proceedings of the*  
30 *Royal Society B*, **265**, 427–31.
- 31 Roberts, G. (2005). Cooperation through interdependence. *Animal Behaviour*, **70**, 901–8.
- 32 Roberts, G. and Renwick, J.S. (2003). The development of cooperative relationships: an experiment.  
33 *Proceedings of the Royal Society B*, **270**, 2279–83.
- 34 Roberts, G. and Sherratt, T.N. (1998). Development of cooperative relationships through increasing  
35 investment. *Nature*, **394**, 175–9.
- 36 Rutte, C. and Taborsky, M. (2007). Generalized reciprocity in rats. *PLoS Biology*, **5**, 1421–5.
- 37 Salmon, C. (1998). The evocative nature of kin terminology in political rhetoric. *Politics and the Life*  
38 *Sciences*, **17**, 51–7.
- 39 Searcy, W.A. and Nowicki, S. (2005). *The evolution of animal communication: reliability and deception in*  
40 *signaling systems*. Princeton University Press, Princeton, NJ.
- 41 Seinen, I. and Schram, A. (2006). Social status and group norms: indirect reciprocity in a helping  
42 experiment. *European Economic Review*, **50**, 581–602.
- 43 Semmann, D., Krambeck, H.-J., and Milinski, M. (2004). Strategic investment in reputation. *Behavioral*  
44 *Ecology and Sociobiology*, **56**, 248–52.
- 45 Shang, J. and Croson, R. (2006). The impact of social comparisons on nonprofit fundraising. *Research in*  
46 *Experimental Economics*, **11**, 143–56.
- 47 Shinada, M. and Yamagishi, T. (2007). Punishing free-riders: direct and indirect promotion of cooperation.  
48 *Evolution and Human Behavior*, **28**, 330–9.



- 1 Simpson, B. and Willer, R. (2008). Altruism and indirect reciprocity: the interaction of person and situation  
2 in prosocial behavior. *Social Psychology Quarterly*, **71**, 37–52.
- 3 Smith, E.A. (2004). Why do good hunters have higher reproductive success? *Human Nature*, **15**, 343–64.
- 4 Smith, E.A. and Bliege Bird, R. (2000). Turtle hunting and tombstone opening: public generosity as costly  
5 signaling. *Evolution and Human Behavior*, **21**, 245–62.
- 6 Smith, E.A., Bliege Bird, R., and Bird, D.W. (2003). The benefits of costly signaling: Meriam turtle hunters.  
7 *Behavioral Ecology*, **14**, 116–26.
- 8 Sober, E. and Wilson, D.S. (1998). *Unto others: the evolution and psychology of unselfish behavior*. Harvard  
9 University Press, Cambridge, MA.
- 10 Soetevent, A.R. (2005). Anonymity in giving in a natural context—a field experiment in  
11 30 churches. *Journal of Public Economics*, **89**, 2301–23.
- 12 Sosis, R. (2004). The adaptive value of religious ritual. *American Scientist*, **92**, 166–72.
- 13 Stephens, D.W. (2000). Cumulative benefit games: achieving cooperation when players discount the future.  
14 *Journal of Theoretical Biology*, **205**, 1–16.
- 15 Stephens, D.W., McLinn, C.M., and Stevens, J.R. (2002). Discounting and reciprocity in an Iterated  
16 Prisoner's Dilemma. *Science*, **298**, 2216–18.
- 17 Stewart-Williams, S. (2007). Altruism among kin vs. nonkin: effects of cost of help and reciprocal exchange.  
18 *Evolution and Human Behavior*, **28**, 193–8.
- 19 Stevens, M. (2005). The role of eyespots as anti-predator mechanisms, principally demonstrated in the  
20 Lepidoptera. *Biological Reviews*, **80**, 573–88.
- 21 Sylwester, K. and Roberts, G. (2010). Cooperators benefit through reputation-based partner choice in  
22 economic games. *Biology Letters*, **6**, 659–62.
- 23 Takahashi, C., Yamagishi, T., Tanida, S., Kiyonari, T., and Kanazawa, S. (2006). Attractiveness and  
24 cooperation in social exchange. *Evolutionary Psychology*, **4**, 315–29.
- 25 Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift fur Tierpsychologie*, **20**, 410–33.
- 26 Tooby, J. and Cosmides, L. (1996). Friendship and the Banker's Paradox: other pathways to the evolution  
27 of adaptations for altruism. *Proceedings of the British Academy*, **88**, 119–43.
- 28 Trivers, R. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, **46**, 35–57.
- 29 Van den Bergh, B. and Dewitte, S. (2006). The robustness of the 'Raise-the-Stakes' strategy: coping with  
30 exploitation in noisy Prisoner's Dilemma games. *Evolution and Human Behavior*, **27**, 19–28.
- 31 Van Lange, P.A.M., Ouwerkerk, J.W., and Tazelaar, M.J.A. (2002). How to overcome the detrimental  
32 effects of noise in social interaction: the benefits of generosity. *Journal of Personality and Social  
33 Psychology*, **82**, 768–80.
- 34 Vogel, G. (2004). The evolution of the Golden Rule. *Science*, **303**, 1128–31.
- 35 Wedekind, C. and Braithwaite, V.A. (2002). The long-term benefits of human generosity in indirect  
36 reciprocity. *Current Biology*, **12**, 1012–15.
- 37 Wedekind, C. and Milinski, M. (2000). Cooperation through image scoring in humans. *Science*, **288**, 850–2.
- 38 West, S.A., Gardner, A., Shuker, D.M., *et al.* (2006). Cooperation and the scale of competition in humans.  
39 *Current Biology*, **16**, 1103–6.
- 40 West, S.A., Griffin, A.S., and Gardner, A. (2007). Social semantics: altruism, cooperation, mutualism,  
41 strong reciprocity and group selection. *Journal of Evolutionary Biology*, **20**, 415–32.
- 42 Williams, G.C. (1966). *Adaptation and natural selection*. Princeton University Press, Princeton, NJ.
- 43 Wilson, M. and Daly, M. (1997). Life expectancy, economic inequality, homicide, and reproductive timing  
44 in Chicago neighbourhoods. *British Medical Journal*, **314**, 1271–4.
- 45 Wilson, M. and Daly, M. (2004). Do pretty women inspire men to discount the future? *Biology Letters*, **271**,  
46 S177–9.
- 47 Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and  
48 Social Psychology*, **51**, 110–16.





- 1 Yamagishi, T. (2003). The group heuristic: a psychological mechanism that creates a self-sustaining system
- 2 of generalized exchanges. *Paper prepared for workshop on 'The Co-evolution of Institutes and Behavior'*,
- 3 Sante Fe Institute, 10–12 January, 2003.
- 4 Yamagishi, T. and Kiyonari, T. (2000). The group as the container of generalized reciprocity. *Social*
- 5 *Psychology Quarterly*, **63**, 116–32.
- 6 Zahavi, A. and Zahavi, A. (1997). *The handicap principle: a missing piece of Darwin's puzzle*. Oxford
- 7 University Press, New York.

