

Reputation

To appear in:
Handbook of Evolutionary Psychology, 2nd Edition (Ed. David Buss)

Pat Barclay
Department of Psychology
University of Guelph
50 Stone Rd. E.
Guelph, ON, Canada, N1G 2W1
Phone: 1-519-824-4120 ext. 58247
Fax: 1-519-837-8629
barclayp@uoguelph.ca

Introduction

Evolutionary scientists are increasingly coming to the conclusion that an organism's reputation can affect its social and reproductive success. Humans are not the only organisms to have reputations: behavioural ecologists now recognize that non-human animals use a combination of observation and personal experience to determine who to challenge, avoid, or mate with (e.g., McGregor & Peake, 2000). Humans rely even more on reputation because language allows us to transmit information to those who do not directly observe events (Smith, 2010). Because of its ubiquity, reputational factors constitute a major selective force in human evolution.

What consequences has this had for the evolution of human behaviour? I will argue that reputation is at least partly responsible for the high levels of non-kin cooperation found in humans, and has also affected the evolution of violence. An organism benefits when others believe that it is willing and capable of conferring benefits and imposing costs on others. Such an organism will be chosen for cooperative interactions and be avoided in competition, both of which historically impacted, and perhaps currently impact, social and reproductive success. This creates competition to be – and be seen to be – a better partner and tougher competitor than others. I will review some of the evidence of how reputations have affected the evolution of human cooperation and conflict. To better understand the impact of reputations on evolution, we need to first establish what reputation is, why it matters, and what sort of information organisms will track about others.

What is reputation?

An organism's reputation in a particular domain is the belief – held by others – that it possesses a particular trait. Reputation is specific to a trait: others believe that you do, or do not, possess that trait. Such traits can be physical (e.g. athletic, strong fighter), dispositional (e.g. honest, faithful, hard-working, willing to escalate fights), social (e.g., has powerful allies), or a combination of these. These reputations for various traits are distinct and potentially separable: for example, one's reputation as a basketball player is different from one's reputation as a brawler, cooperator, hard worker, liar, alcoholic, or lover. The same organism could be high on one of these traits and low on others. There can be overlap between traits – one's reputation in one trait may generalize to conceptually related domains, especially when possession of one trait predicts possession of another. For example, if honesty and cooperative intent are both caused by the same underlying psychology, then one's reputation for honesty will affect others' beliefs of one's willingness to cooperate with others. The degree of generalization should depend on how well one trait predicts another. An overall "good reputation" implies that most others view a person positively on a number of relevant traits.

One's reputation is not absolute or objective: it exists solely in the minds of others. Each individual must form its own impression of everyone else on various traits, using a combination of personal experience, observation, physical or behavioural cues, and information transmitted from others (gossip). These impressions may be accurate or inaccurate, and impressions may vary from person to person due to misperceptions, biases, or different interaction histories. For example, my ally may be perceived as honest by my coalition members, yet other coalitions may

perceive him as dishonest – the other coalitions may be biased, may have misinterpreted his actions, or perhaps my ally actually *is* less honest when dealing with rival coalitions.

Thus, in its simplest definition, “reputation” is a simple function of others’ beliefs, e.g. the average belief held by relevant audiences. More complex definitions may rely on a perception of what others think, i.e. a belief about how other people view someone. This more complex definition limits the study of reputation by restricting it to species with a Theory of Mind, situations with multiple observers who all have the opportunity to assess others’ beliefs, and cases where audience members generally agree. Because of these limits, the simplest definition of reputation is preferable because it is more general.

Why does reputation exist?

It is obviously advantageous to remember what others have done to you: this allows you to assess their likelihood of doing it again. You can then approach those who are likely to confer benefits upon you in the future, and avoid those who will impose costs. However, direct interaction carries potential costs such as losing a fight or being cheated. It pays to predict what others will do before directly interacting with them, for example by observing them interact with third-parties (Dabelsteen, 2005).

Many studies show that non-human animals “eavesdrop” on the interactions or communications of others in order to gain useful information (McGregor & Peake, 2000). For example, male and female Siamese fighting fish (*Betta splendens*) assess other males’ fighting ability by watching them fight, and then approach or avoid them as appropriate (Doutrelant & McGregor, 2000; McGregor & Peake, 2000; Oliveira et al., 1998). Female great tits (*Parus major*) listen to the outcomes of male-male interactions and preferentially approach winners to assess them for extra-pair copulations (Otter et al., 1999, 2001). Sexually experienced female Japanese quail (*Coturnix japonica*) avoid males they have seen being too aggressive (Ophir et al., 2005). Reef fish observe the interactions between cleaner fish and other clients to determine whether to associate with that cleaner (Bshary, 2002; Bshary & Grutter, 2006). These are but a few of the examples showing that non-humans in taxonomically diverse species can observe others’ interactions – in competitive or cooperative situations – to glean important information about the costs and benefits of associating or competing with those others. These examples also show that observing an encounter changes the subsequent behaviour of the observer, which affects the fitness of the individual being observed.

Humans extend the reliance on observation by incorporating the observations of others. Because of language, we can hear about others’ past behaviours (via gossip), and then use that information to assess their ability and willingness to confer benefits or impose costs on us. Such socially-transmitted reputations are often what people mean by “reputation”, but this is simply an extension of the more general case of predicting others’ behaviour based on their interactions with third parties. Hearing about past behaviours gives listeners access to events they did not directly observe. Most conversations are indeed about social topics (Dunbar, 2006), and of those topics, most discussion involved exchanging information on the speakers’ or others’ behaviour and experiences (Dunbar et al., 1997). The use of language allows reputation to be even more effective at shaping behaviour than direct observation alone (Dunbar, 2006; Smith, 2010); the

ability to spread information will effectively increase the size of the “audience”, and thus the fitness consequences of behaviour (Nowak & Sigmund, 2005).

There are time, energetic, and cognitive costs associated with attending to others’ interactions (Peake, 2005); learning by observation requires cognitive abilities beyond the ability to learn from personal reinforcement and punishment. Information may be transmitted inaccurately, like in the children’s game of “Broken Telephone”, or even deliberately manipulated by others for their own gain. An organism may behave differently depending on its partner – X’s interactions with Y are an imperfect cue of how X will interact with Z (Krasnow et al., 2012). Nevertheless, as long as the information carries some statistically predictive ability, it can be beneficial to observe third-party interactions or use socially-transmitted information about others’ past actions, abilities, and general behavioural tendencies.

Influencing one’s reputation and the reputations of others

Given that others are influenced by what they see, or what they hear from others, it pays for an organism to influence how it is viewed by others – to “manage” its reputation. This does not require the ability to attribute mental states to others (Theory of Mind), it simply requires the ability to recognize the presence of an audience, and an evolved or learned decision rule to behave differently when observed.

Many species alter their behaviours when they are watched by a relevant observer. For example: male Siamese fighting fish vary their aggressive displays depending on whether an audience is present and whether the audience is male or female; male vervet monkeys are more affiliative towards infants when the infant’s mother is watching; and male budgerigars spend more time courting extra-pair partners when their primary mate is not watching (all reviewed by Matos & Schlupp, 2005). Cleaner fish give better cleaning service (e.g. fewer bites) to their clients when they are observed by another potential client (Bshary & Grutter, 2006), especially when the observer is a highly desirable client (Bshary, 2002). These cleaner fish may even be deceptive, in that they lure in desirable clients by behaving nicely when observed, only to exploit those desirable clients (Bshary, 2002). Victory displays occur after winning a fight in species ranging from crickets to frogs, from songbirds to canids, and may function to broadcast one’s success to audiences (Bower, 2005). Primates who lose a fight often redirect their aggression towards lower-ranking group members, which may function to display that they are still formidable despite losing to a high-ranking individual (Kazim & Aureli, 2005). As for humans, the effects of observation are so ubiquitous that whole areas of psychology are dedicated to understanding the effects of observation (e.g. social facilitation, impression management), and researchers have to be very careful about how observation may affect the results of their studies (e.g. demand effects, socially desirable responding).

Although impression management does not require Theory of Mind, a Theory of Mind adds considerable strategic complexity by allowing an organism to tailor its impression management according to what an observer knows. Highly social organisms should have psychological mechanisms for monitoring not only the likelihood of being observed, but also the characteristics of the potential audience, the value of a reputation to that audience, how any given act will

change how that audience will see oneself, and how one might even avoid an audience to prevent them from observing one's undesirable behaviours (Barclay, 2013).

Reputations often involve an implicit comparison with others: one is seen as *stronger*, *tougher*, *nicer*, or *more* cooperative than others. As such, organisms should also have psychological mechanisms for monitoring their reputation relative to others and acting accordingly (Barclay, 2013). For example, are one's competitors seen as more generous, and if so, should one compete by acting more generously or by attacking the reputation of the competitors?

Just as it is advantageous to manipulate one's own reputation, it is also advantageous to manipulate audience perceptions of one's allies and competitors (Hess & Hagen, 2006). Gossip is arguably all about influencing the reputations of other people, making one's allies seem better than they are and one's competitors seem worse. We should predict that people will be most likely to spread information about domains that are most important in making allies look good and competitors look bad (e.g. Buss & Dedden, 1990). Given the potential for its manipulation, people assess the veracity of gossip using cues such as the number of sources they hear it from and the vested interests of the person from whom they hear the gossip (Hess & Hagen, 2006; Sommerfeld et al., 2008).

Types of Reputation

The costs and benefits of social interactions depend on who one is interacting with, and the type of interaction, such as a cooperative versus competitive situation. Some individuals are highly capable of conferring benefits upon others (e.g. good hunters), or are more likely to do so (e.g. honest cooperators), whereas other individuals are less willing and able to confer benefits. Some individuals will continue to confer such benefits (e.g. faithful partners), whereas others will not. Some individuals are more capable of imposing costs on others (e.g. good fighters) or are more likely to do so (e.g. chronically angry people), whereas others are less willing and able to impose costs. This is obviously important information to track.

Any organism gains from seeking out situations where it receives benefits from others. Good cooperators are often worth approaching, bad cooperators are usually not. Faithful partners will provide benefits for longer than unfaithful or fraudulent partners. Because others differ in their ability and willingness to confer benefits in both the short and long term, we should expect organisms to track who is most able, willing, and available to do so.

Much evidence shows that humans do judge others based on these three qualities: *abilities*, *tendencies*, and *availability* (reviewed by Barclay, 2013). The value of a cooperative partner is some function of these three traits. The best cooperative partners are very able to help, willing to help, and available to do so. The worst partners are none of those three. Intermediate partners have intermediate levels of these traits, or are high on one but low on others (e.g. able to help but less willing to do so). We should then expect organisms to track this information, and to approach and preferentially help partners who have a reputation for being able to help, willing to help, and available as a cooperative partner. This does not require conscious tracking of these traits or any awareness that they affect one's partner preferences, just as people do not

consciously track the MHC genotypes of their romantic partners (e.g. Alvergne & Lummaa, 2009). Instead, our proximate psychological mechanisms (e.g. emotions) do this tracking for us.

Organism should also avoid situations where others will impose costs that outweigh the benefits. For example, conflict over resources can be worthwhile if one will win the competition, but is usually not worthwhile if one will lose. Courting the spouse of a powerful individual carries high risks, whereas there are fewer costs associated with courting the spouse of a weaker individual or someone who is absent and unable to retaliate. Because others differ in their ability and willingness to impose costs, we should expect organisms to track who is most able, willing, and available to impose costs, and avoid conflict with those who score high on those traits. (Daly & Wilson, 1988; Sell et al., 2009)

This chapter focuses on reputation for cooperation (conferring benefits), and to a certain extent, a reputation for aggression (imposing costs). There are specific examples of these which are beyond the scope of the chapter, for example people carry reputations for commitment and fidelity to their mates or allies. Ultimately, these are specific instantiations of the more general principles of benefit conferral and cost imposition, e.g. a mate's reputation for infidelity indicates there are fewer long-term benefits of associating with him or her. Many of the same principles underlie partner choice for mating and partner choice for other social relationships (Barclay, 2013). As such, many principles that apply to reputations in one domain will also apply to reputations in other domains.

Reputation for Cooperation

Indirect Reciprocity

Axelrod's (1984) seminal computer simulations of the evolution of cooperation in the Prisoners Dilemma game, and the success of conditional cooperators like the Tit for Tat strategy, is an example of *direct reciprocity*: individuals help those who have helped them in the past, or likely will in the future. A conditional cooperator helps those who help, thus reaping the long-term rewards of mutual cooperation, but refuses to be suckered for long by non-cooperators. Years of mathematical models and computer simulations show that most successful strategies involve some conditional willingness to reciprocate help, and much evidence shows that people are more likely to help those who have previously helped them.

Humans go beyond direct reciprocity by also helping others who have not personally helped them, or who will not have an opportunity to reciprocate. Helpful acts may be reciprocated not just by the recipient, but by others who observe it or hear about it; this is known as *indirect reciprocity* (Alexander, 1987; reviewed by Nowak & Sigmund, 2005). Indirect reciprocity works when those who help gain a good reputation and are thus more likely to be helped by observers. Those who refuse to help get a bad reputation and are more likely to be refused help.

Wedekind and Milinski (2000) gave participants the chance to donate to others and have these donations (or lack thereof) be made public to others. Participants were more likely to give to people who had given to others in the past, even though the design ensured that no one would have the opportunity to reciprocate a donation directly. Subsequent experiments have also shown

that people give more to generous people (Milinski et al., 2001; Wedekind & Braithwaite, 2002; Seinen & Schram, 2006; Semmann et al., 2004). People base their giving on a combination of personal experience and social information about others (Roberts, 2008; Sommerfeld et al., 2007); the more positive things we hear about someone, the more likely we are to give to that person (Sommerfeld et al., 2008). Field research suggests that people gossip about the cooperation of others and that this has “real economic consequences” (Fessler, 2002; Kniffin & Wilson, 2005). For example, hunters who share meat are more likely to receive meat from group members (Gurven et al., 2001), though this could also be because group members have a vested interest in the well-being of food providers (Barclay & Van Vugt, in press).

Given that people pay attention to others’ cooperativeness, it pays to be more cooperative when others are watching. Much research shows that the presence of observers increases “good behaviour” in many domains, including donations in monetary games within laboratories (e.g. Barclay, 2004; Hardy & Van Vugt, 2006; Hoffman et al., 1994; Milinski, 2002; Rege & Telle, 2004), willingness to volunteer (Bereczkei et al., 2007), contributions towards educating others about climate change (Milinski et al., 2006), voter turnout (Gerber et al., 2008), and simulated tax donations (Coricelli et al., 2010). People are also more likely to cooperate if others might gossip about them (Feinberg et al., 2014; Piazza & Bering, 2008). This effect of observation can be harnessed to promote cooperation in many situations including contributions to public goods (Milinski et al., 2002; Panchanathan & Boyd, 2004), responsible consumerism (Barclay, 2012; Griskevicius et al., 2012), and the fight against climate change (Milinski et al., 2006). Increased giving under observation is sometimes strategic (Semmann et al., 2004; Barclay & Willer, 2007), but it is also possible that cooperative emotions like empathy and guilt are experienced more strongly in the presence of observation – this requires future investigation.

There are different types of indirect reciprocity that differ in what constitutes a “good” (or “bad”) act that is worthy (or unworthy) of reciprocation (reviewed by Nowak & Sigmund, 2005). In some models, helping anyone is seen as good and increases one’s reputation (“image scoring”). In other models (“standing strategies”), helping a defector does not increase one’s reputation, and it may even be seen as “bad” to help someone who is “unworthy” of help. “Image scoring” is less likely to be evolutionarily stable than variants of “standing strategies”, because in the former it would not pay to discriminate against defectors (Ohtsuki & Iwasa, 2004, 2006, 2007). However, current experimental evidence suggests that people use image scoring rather than standing strategies (Milinski et al., 2001). This may be because it is difficult to tell whether a defection against a defector is truly “justified” (see also Barclay, 2006), and whether such defection represents moralistic discrimination or a cheap excuse to cheat someone—a topic for future work.

Signals of Ability to Confer Benefits

Some acts are difficult to perform, and can be done best by those with special abilities. For example, it takes strength and agility to leap from a moving boat to catch a 150 kg green sea turtle; uncoordinated individuals would be less likely to catch turtles. Billionaires can give away sums of money that would bankrupt normal people. Good swimmers can dive into raging rivers to save drowning babies, whereas bad swimmers might drown. Because these acts are easier or less costly for some people to perform, they carry information about the performer: the ability to

share turtle meat at feasts is a credible signal of the hunter's strength and agility (Smith & Bliege Bird, 2000), Bill Gate's billion dollar donations are a credible signal of his vast wealth, and diving into a river to save a baby is a credible signal of swimming ability. Such acts thus convey information about an individual's agility, strength, wealth, and other such qualities, all of which are desirable in social partners because they indicate an ability to confer benefits on others.

These are all examples of costly signaling theory (Grafen, 1990; Zahavi, 1977), where the cost of a signal is used to maintain signal honesty. It would be beneficial for anyone to appear strong, agile, and wealthy, but the fitness costs of some acts are not worth it for someone who does not actually possess the necessary qualities (Gintis et al., 2001; Searcy & Nowicki, 2007). For example, even if I managed to borrow one billion dollars to give to charity, the cost of bankruptcy would far outweigh any reputational benefits to me. Such a donation would be worth it for Bill Gates, because for him the cost is trivial. Similar arguments hold for the other examples: the anticipated hunting success (and subsequent reputation) is worth the time and risk for a good hunter but not a bad hunter, and a reputation as a baby-saving hero is only worth the drowning risk for someone who is unlikely to drown. Thus, the cost (or potential cost; Getty, 2006) deters those who do not possess the necessary qualities, so they can be used as honest signals of one's qualities. Audiences benefit from attending to those signals and gaining useful information. Again, this does not imply that people consciously assess the costs and benefits or consciously track them in others; our emotions (e.g. fear, bravado, empathy) do this for us.

Costly signaling theory has been used to explain many types of extravagant helping, including lavish sharing at feasts (Boone, 1998), large-scale philanthropy (Harbaugh, 1998), big-game hunting (Hawkes & Bliege Bird, 2002; Smith & Bliege Bird, 2000), and blood donations (Lyle et al., 2009). These all require wealth, political connections, physical abilities, or health, which are all related to the actor's ability to confer benefits on others. There may be direct reproductive advantages for such behaviour: for example, good hunters have more children than poor hunters (Smith, 2004), including more children with other men's wives (Hill & Kaplan, 1988). Of course, the benefits of such signaling need not be in terms of mate attraction – those who possess such abilities may be chosen more often as allies or avoided more often as competitors (Smith & Bliege Bird, 2000).

The above examples all involve conspicuous generosity as a signal of resources or abilities, but obviously many signals do not involve generosity. Conspicuous consumption and conspicuous leisure have been seen as signals of wealth for over 100 years (Veblen, 1899). Physical abilities could be signaled via athletic displays and sporting wins, intellectual abilities could be signaled via wit, vocabulary, or problem-solving, and so on (Barclay, 2013). In fact, signaling via non-generous means is arguably more common than signaling via generosity. Signaling via generosity may also carry information about one's character, which is beneficial, but there may be a risk of "diluting the signal" by signaling more than one trait in a single act. Future research should investigate when people will signal their traits via generosity, and whether this is as effective as using non-generous means like conspicuous consumption.

Signals of Willingness to Confer Benefits

If someone has helped you in the past, it suggests that they are more likely to help in the future as well (André, 2010). This generalizes beyond established pairs: people who are cooperative within one group tend to be cooperative within other groups also (Kurzban & Houser, 2005). This is the basis of stable personality traits like agreeableness: niceness generalizes across situations. Someone who creates a reputation for helping others is essentially broadcasting their willingness to confer benefits on others. Similarly, someone with a reputation for commitment has succeeded in broadcasting his or her willingness to provide benefits to partners (McNamara & Houston, 2002).

Why maintains the honesty of such signals? The previous section described how signals of abilities are kept honest by the high potential cost of extravagant generosity (Gintis et al., 2001). By contrast, many signals of willingness to help do not appear very costly and could seemingly be done by anyone. It does not require wealth or athletic ability to spend time with someone, groom them, or to volunteer in a soup kitchen. Such acts cost the same time for anyone. In these cases, honesty is maintained not by differential costs, but by differential benefits: it would not be worth it to cooperate at time A if one intended to cheat at time B and lose out on future cooperative interactions (André, 2010; Bolle 2001; Ohtsubo & Watanabe, 2009; Smith & Bliege Bird, 2005). Signals will be honest as long as the cost of public helping is: a) greater than the immediate benefits of “suckering” someone; and also b) less than the long-term benefits of mutual cooperation. The former condition makes cheaters not bother trying to appear cooperative, whereas the latter condition makes it pay off for long-term cooperators to broadcast their willingness to help. Thus, honesty is maintained by differential long-term benefits accrued to cheaters and cooperators, not differential costs (Barclay & Reeve, 2012; Grafen, 1990).

Not surprisingly, people treat public helping as though it carries information about the helper’s future trustworthiness. People entrust more money to those who have given money to a charity or public good (Albert et al., 2007; Barclay, 2004, 2006; Keser, 2003), preferentially associate with those who have given to others (Barclay & Willer, 2007; Feinberg et al., 2014), and prefer generous people to neutral controls for romantic relationships (Barclay, 2010). This adds up to substantial benefits for those with a good reputation. In business, online reputation systems are prevalent in places like e-Bay, Amazon, and TripAdvisor; these are designed so that sellers can acquire a good reputation, and this can directly benefit honest companies (Frank, 2004). Cooperative reputations are so valuable that they are worth maintaining in order to sell a reputable business (Pfeiffer et al., 2012), or possibly to pass along to offspring.

Public helping: indirect reciprocity or costly signal of cooperative intent?

There are many overlapping predictions if we view helping behaviour as indirect reciprocity versus as a costly signal of cooperative intent. For example, both theories predict that organisms will be more cooperative while observed, be concerned about their reputation, attempt to enhance their reputation, and so on. In fact, these two theories may not even be separate: reciprocation itself could be seen as a signal of future willingness to help (André, 2010). Indirect reciprocity may simply be the outcome of organisms attempting to assess the probability that another organism will cooperate in the future, combined with their tendency to signal their own

willingness to cooperate. A similar argument has been made about moral judgment: when people judge the morality of acts, perhaps what they are really assessing is the probability that the actor is a good person and future cooperator (Pizarro et al., 2012; Tannenbaum et al., 2011).

Some evolutionary psychologists (e.g. McCullough et al., 2013; Sell et al., 2009) argue that we observe others' actions in order to assess their "Welfare Trade-off Ratio" (WTR) towards us, i.e. how much they value our welfare relative to their own. Some acts imply a high WTR (i.e., actor values our welfare), some imply low WTR (i.e., actor does not value us), and some even imply a negative WTR (i.e., actor values our demise). What currently appears to be indirect reciprocity could simply be people attempting to assess the welfare trade-off ratio of others and then initiate or maintain positive relationships with those who appear likely to cooperate in the future. Future theoreticians and empiricists should test whether indirect reciprocity is simply the outcome of this same process, with different acts having different predictive ability of one's future cooperation.

Competitive Helping

Individuals differ in their ability and willingness to confer benefits on social partners like allies, friends, and mates. Whenever organisms can choose whom to interact with, this creates a market-like competition over the "best" partners (Noë & Hammerstein, 1994, 1995). The best way to attract a good partner is to be a good partner, so each organism gains from appearing more able or more willing to confer benefits on its partners. Much social competition is about demonstrating abilities (e.g. sports), but one can also compete using generosity: there is a market-based incentive to compete to be more generous than others, in order to attract more social partners and/or higher quality partners. This process of "competitive altruism" or "competitive helping" consists of not just appearing nice, but appearing *nicer* than one's competitors (Barclay, 2004, 2011, 2013; Barclay & Willer, 2007; Roberts, 1998; Van Vugt et al., 2007).

Competing over social partners is similar to competing over mates, and many of the same principles apply to both (Barclay, 2013). In fact, sexual selection is just a specific instantiation of social selection, which is when one's fitness depends on the actions and choices of others (West-Eberhard, 1979, 1983). In this particular case, individuals compete over non-romantic relationships just as they do over romantic relationships. This competition over social partners can lead to a "runaway" process towards higher levels of generosity (McNamara et al., 2008; Nesse, 2007), up to the point where the marginal benefits of attracting additional partners is outweighed by the marginal costs spent to attract them (Barclay, 2011, 2013).

Experimental evidence shows that people actively escalate their generosity when it can affect others' choice of partners. For example, people give more money in laboratory experiments when observed by others (e.g. Hardy & Van Vugt, 2006; Rege & Telle, 2004), but they give the most when those observers can choose whom to interact with in the future (Barclay, 2004; Barclay & Willer, 2007; Sylwester & Roberts, 2010). These latter findings show that people are not only trying to appear nice, but are actively trying to appear *nicer* than competitors. People will also compete to give more to environmental charities when it will affect observers' choice of

partners, and this effect is above and beyond the effect of simply being observed (Barclay & Barker, in preparation).

Additionally, other research shows that generous people are accorded higher status, both in laboratory tasks (Hardy & Van Vugt, 2006; Willer, 2009) and in anthropological studies (Price, 2003). Extravagant helping may be a way to compete over mates and social partners: anthropological examples such as big-game hunting, large-scale philanthropy, and hosting large feasts have all been interpreted as competition to be more generous than others (Barclay, 2013; Boone, 1998; Harbaugh, 1998; Hawkes & Bliege Bird, 2002; Smith & Bliege Bird, 2000).

Reputation for Aggression

Ability and Willingness to Inflict Costs

Humans excel at non-kin cooperation, but also use aggression. Fatal and non-fatal conflict is endemic in non-state societies, with many men dying violently at the ends of other men (Chagnon, 1997; Daly & Wilson, 1988; Pinker, 2011; Puts, 2010). Given the prevalence and costs of violence, it obviously pays to know who to avoid challenging. Organisms can assess this from personal experience, valid cues like physical size or behavioural displays (Sell et al., 2009), observations about others' willingness to engage in conflict or risk-taking in general (Fessler et al., 2014; Johnstone & Bshary, 2004), observations of the outcomes of others' fights (McGregor & Peake, 2000), or indirectly hearing about any of the above, i.e. transmitted reputation.

In their seminal book on human aggression, *Homicide*, Martin Daly and Margo Wilson (1988) summarize the importance of a formidable reputation in many environments, i.e. a reputation for being willing and able to inflict costs on others in response to affronts. A complete summary of their evidence is beyond the scope of this chapter, but it is worth re-quoting their oft-quoted words:

Men are known by their fellows as “the sort who can be pushed around” or “the sort who won't take any shit,” as people whose word means action and people who are full of hot air, as guys whose girlfriends you can chat up with impunity or guys you don't want to mess with (Daly & Wilson, 1988, p. 128).

Exactly what information is assessed and passed on about others' ability to inflict costs? Physical abilities, intellectual abilities, and political connections can all be used to help or hurt someone. Many or most traits discussed in the section on reputations for “Ability to Confer Benefits” would also apply to one's ability to inflict costs – the same reputation is useful for both attracting allies and deterring competitors. For example, sporting ability can signal one's ability to physically confer benefits or physically impose costs. Future work should determine the relative importance of these two abilities – benefit-conferral and cost-imposition – for people's reputations, in order to determine when and why audiences attend to certain signals.

Some displays may be designed to enhance or repair one's reputation for formidability. Many animals perform victory displays after winning a fight, which can broadcast their success – and corresponding formidability – to others who may not have observed the victory (Bower, 2005);

this display function has also been suggested for the postures of human athletes after a triumph (Matsumoto & Hwang, 2012). Many primates show redirected aggression, where the loser of a fight aggresses against someone else even lower in the hierarchy, which could potentially function to deter challenges from others by signaling the loser's residual formidability (Kazim & Aureli, 2005). Even the willingness to fight itself could signal one's formidability, because fighting is less costly for more formidable individuals such that will engage in it more readily (Benard, 2013; Johnstone & Bshary, 2004).

People's aggression is certainly affected by opportunities for reputation (reviewed by Benard, 2013; Daly & Wilson, 1988; Felson, 1978; Frank, 1988). For example, men are more likely to violently retaliate against transgressions if there is an audience than when there is no audience (reviewed by Felson, 1978). This should be unsurprising to anyone who has observed a physical fight in school, in a bar, or elsewhere. Laboratory experiments confirm that opportunities for reputation cause people to challenge others more often over resources in an attempt to convey high competitive ability (Benard, 2013). Furthermore, much research shows that people are more likely to back down from aggressive confrontations if they can do so without "losing face", i.e. without gaining a reputation for cowardice (reviewed by Daly & Wilson, 1988; Felson, 1978). In laboratory experiments, status motives make men more likely to engage in face-to-face confrontations and women's indirect aggression (Griskevicius et al., 2009). Reputation is also involved in bargaining, with people attempting to establish a reputation as a "tough bargainer" in order to receive better bargains in the future, even if it means engaging in irrational behaviour in the present (DeClerck et al., 2009; Frank, 1988; Nowak et al., 2000; Yamagishi et al., 2009).

A reputation for toughness may seem at odds with a reputation for cooperation: if people value cooperation, wouldn't they avoid highly aggressive individuals? These two qualities – conferring benefits on others versus imposing costs - have different values in different environments. In environments characterized by social exchange and with central authorities to limit interpersonal conflict, the former will be more important for social success. In environments with intense competition over limited resources, the latter will be more important. When people form alliances to aggressively compete with other alliances, then both are important. Ultimately, the best partners are those who are highly able to confer benefits and impose costs, and are highly willing to confer benefits specifically upon oneself and impose costs specifically on one's rivals.

Harnessing the Power of Reputation

Given that people are so concerned about reputations, we can use this knowledge to promote prosocial behaviours and decrease antisocial behaviours (Barclay, 2012). For example, people who are made to think about status and good reputation tend to make more benevolent decisions (Griskevicius et al., 2007) and purchase more environmentally-friendly products (Griskevicius et al., 2010). People do more to preserve the environment when observed than when anonymous (Milinski et al., 2006), and will even compete to give more to the environment (Barclay & Barker, in preparation). After being told about others' high cooperation, people are more likely to give to fundraisers (Shang & Croson, 2006), cut energy use (Allcott, 2011), and reuse hotel towels Goldstein et al., 2007). Under some circumstances it could even be useful to limit reputational opportunities, for example to reduce aggressive retaliations and escalations of conflict.

Even false cues of reputation can be effective at changing behaviour. Observation is one component of reputation, and photographs of eyes (a false cue of observation) have been shown to trigger higher monetary donations in laboratory games (Burnham & Hare, 2007; Haley & Fessler, 2005; Mifune et al., 2010), more payment for coffee on an “honour system” (Bateson et al., 2006), more cleaning of litter (Ernest-Jones et al., 2011; Francey & Bergmuller, 2012), higher charitable donations in supermarkets (Ekström, 2011), and lower bike thefts (Nettle, 2012). However, the effects of such false cues may be transient: people quickly habituate to images of eyes (Sparks & Barclay, 2013), and eventually come to ignore verbal punishment that is not followed by tangible consequences (Sparks & Barclay, in preparation). Would-be social engineers would be unwise to rely forever on false cues of reputation, unless those cues are at least occasionally followed by real opportunities for reputational costs and benefits.

Despite the benefits of harnessing reputation, there are several risks associated with doing so. Barclay (2012) identified the following limitations and unknowns, in increasing order of importance: 1) reputational benefits must outweigh the costs of helping; 2) reputational cues must be stronger than other situational factors; 3) people may habituate to non-informative cues of reputation; 4) not everyone values reputation; 5) reputation only pays off in the long-term; 6) extrinsic incentives may “crowd out” intrinsic motivations; 7) reputation can promote negative behaviours like aggression; 8) reputations can be manipulated; and 9) publicly identifying reputational incentives may reduce the benefits to cooperators and thus undermine cooperation. It is important to understand and overcome these limitations before relying on untested means of harnessing reputation.

Future directions: towards a more comprehensive science of reputation

Evolutionary researchers have learned much about the power of reputation and how it has affected the evolution of cooperation and conflict. Despite these advances, there are currently many unknowns about the evolution and dynamics of reputation. The following are some future directions that warrant investigation:

Broader roles of reputation: To what extent does reputation underlie other phenomena? For example, in his classic book *Passions Within Reason*, Robert Frank (1988) ultimately relies on reputation as the reason why emotions are hard-to-fake signals of future intent. Ohtsubo and Watanabe (2009) argue that the costs of apologies make them effective as signals of cooperative intent. Other researchers rely on reputation when they argue that religious rituals serve as a costly signal of cooperative intent towards fellow believers (Sosis, 2004). Similarly, could xenophobia be a signal to in-group members that one is committed to cooperate with (and only with) other in-group members, with the honesty of the signal maintained by the opportunity costs of foregone partnerships with out-group members? Are moral judgments a way of advertising one’s beliefs – and thus future behaviour – to audience members? What other phenomena might ultimately rely on reputation?

Proximate mechanisms: What proximate psychological mechanisms have been selected for as a result of past reputational consequences? To what extent has this resulted in a genuine concern for others (e.g. see Barclay, 2013) versus simply a conscious concern for reputation? Are

reputational effects caused by an increase in the causal emotions themselves, e.g. does genuine empathy (cooperation) or anger (aggression) increase in response to the presence of an audience?

Interactions between different reputations: How do different types of reputation interact, such as a reputation for conferring benefits versus imposing costs, or a reputation for ability versus willingness to confer benefits? How and why does a reputation for one trait affect one's reputation for other traits? If one act signals multiple traits, is there a risk of "diluting" the signal across too many domains? What is the optimal balance between an able partner versus a willing partner, or a partner who both confers benefits and imposes costs, and how does this affect what information people track and transmit about others?

Getting into specifics: What traits are signaled by what acts? How useful are different acts at conveying information about the actor, and how much do audiences rely on them? Is this information passively conveyed as a byproduct of the actor's normal actions ("cues"), or is the information actively transmitted and exaggerated by an actor which has evolved to perform that action for its information value ("signals")?

Information value of different actions: What affects the honesty of the signals that affect one's reputation? For example, exactly why does cooperation at time A predict cooperation at time B, i.e. why do stable individual differences exist? If the honesty of signals is maintained by costs, then what types of costs are involved (e.g. performance costs vs. opportunity costs, Barclay & Reeve, 2012). There is much theoretical work on costly signals of stable traits like genetic quality, but much less done on signals of intent or future behaviour.

Novel environments, plasticity, and the importance of reputation: How does reputation in today's world differ from reputation in ancestral environments? What effect does this have? For example, given that most of us no longer live in small tight-knit societies where everyone knows everyone's business, does this diminish the importance of reputation? To what extent can people adjust to the changing role of reputation, or are our evolved reputation-based emotions no longer as adaptive as they once might have been (Barclay & Van Vugt, in press)? Will the internet compensate for this? Which false cues of reputation will people readily habituate to (e.g. photos of eyes, Sparks & Barclay, 2013), and which will continue to have an effect?

Novel environments, plasticity, and evolved cues of underlying traits: Are some cues less informative in modern environments than in ancestral environments, and how do people react to those? For example, politicians' emotional rhetoric is arguably a less reliable signal of their cooperative intent than would be the case in a small reputation-based band, yet people still seem to treat it as a valid cue. To what extent will people continue to rely on ancestral cues versus show adaptive plasticity in what cues they rely on?

Dealing with new actions, cues, and signals: How does a given act initially come to signal a given trait, such that people then track those acts in others' reputations (Panchanathan & Boyd, 2004)? For example, how could an act like protecting the environment come to signal good character or to be valued within a system of indirect reciprocity (Barclay, 2012)? Does this require a pre-existing correlation between character traits and a given act (e.g. nice people just

happen to support the environment), which audiences pick up on, and which is then later exaggerated by the actors as an active signal?

Audience skepticism: Given that people tend to behave differently when observed, how does this affect the information that can be inferred from someone's public actions? To what extent do audiences change their impressions of someone's actions depending on the number and nature of other audience members, e.g. how skeptical should one be of public generosity relative to private generosity? What happens when people become aware of the reputational consequences of various actions? For example, will people trust cooperators less if they know that cooperators can benefit from their actions (Barclay, 2012)? If so, this leads to a recursive problem, because it would affect what level of cooperation would be displayed, which then affects skepticism, and so on in a feedback cycle... how can we resolve this?

These are just a few of the questions that remain when attempting to understand reputation. The science of reputation is just getting started, so we should look forward to more theoretical and empirical investigations of these questions. Eventually we should hope to see predictions that are much more nuanced than "people will be nicer when observed", and be able to quantify exactly how much nicer, when, in what situations, to whom, and exactly how audiences will respond.

Conclusions

Reputations are an important part of the lives of many social organisms. Humans' ability to transmit information verbally has increased the importance of reputation because it gives individuals access to information they did not directly observe. Each individual's reputation affects how others act towards it, such that reputations have real fitness consequences. An organism benefits from being seen as more effective at conferring benefits on allies and imposing costs on competitors. This selects for higher levels of cooperation, but also higher levels of aggression - and manipulation of the appearance of both - as organisms compete to have a better reputation than others. This can occur in any organism, but is particularly relevant in humans because language makes reputations much more important for us than for other species. As a result of such past selective pressures, humans most likely have psychological adaptations specifically for tracking the reputations of others, monitoring their own reputations, adjusting their behaviour according to the reputational consequences, and manipulating information to make themselves look better and rivals look worse. By understanding the role of reputation in our daily lives, and its role in the evolution of human behaviour, we can be more effective at harnessing its power to promote positive change.

References

- Albert, M., Güth, W., Kirchler, E., & Maciejovsky, B. (2007). Are we nice(r) to nice(r) people? An experimental analysis. *Experimental Economics*, *10*, 53-69.
- Alexander, R. D. (1987). *The Biology of Moral Systems*. New York, NY: Aldine de Gruyter.
- Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, *95*, 1082-1095.
- Alvergne, A., & Lummaa, V. (2009). Does the contraceptive pill alter mate choice in humans? *Trends in Ecology and Evolution*, *25*, 171-179.
- André, J.-B. (2010). The evolution of reciprocity: social types or social incentives? *The American Naturalist*, *175*(2), 197-210.
- Axelrod, R. (1984). *The Evolution of Cooperation*. New York, NY: Basic Books.
- Barclay, P. (2004). Trustworthiness and Competitive Altruism Can Also Solve the “Tragedy of the Commons”. *Evolution & Human Behavior*, *25*, 209-220.
- Barclay, P. (2006). Reputational benefits for altruistic punishment. *Evolution and Human Behavior*, *27*, 325-344.
- Barclay, P. (2010). Altruism as a courtship display: Some effects of third-party generosity on audience perceptions. *British Journal of Psychology*, *101*, 123-135.
- Barclay, P. (2011). Competitive helping increases with the size of biological markets and invades defection. *Journal of Theoretical Biology*, *281*, 47-55.
- Barclay, P. (2012). Harnessing the power of reputation: strengths and limits for promoting cooperative behaviours. *Evolutionary Psychology*, *10*(5), 868-883.
- Barclay, P. (2013). Strategies for cooperation in biological markets, especially for humans. *Evolution & Human Behavior*, *34*(3), 164-175.
- Barclay, P., & Barker, J. (in preparation). Greener than thou: competitive helping results in higher contributions to environmental charities.
- Barclay, P., & Reeve, H.K. (2012). The varying relationship between helping and individual quality. *Behavioral Ecology*, *23*(4), 693-698.
- Barclay, P., & Van Vugt, M. (in press). The evolutionary psychology of human prosociality: adaptations, mistakes, and byproducts. To appear in D. Schroeder & W. Graziano (Eds.) *Oxford Handbook of Prosocial Behavior*. Oxford, UK: Oxford University Press.

- Barclay, P., & Willer, R. (2007). Partner choice creates competitive altruism in humans. *Proceedings of the Royal Society of London Series B*, 274, 749-753.
- Bateson, M., Nettle, D. & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters*, 2, 412-414. doi:10.1098/rsbl.2006.0509
- Benard, S. (2013). Reputation systems, aggression, and deterrence in social interaction. *Social Science Research*, 42, 230-245.
- Berezkei, T., Birkas, B., & Kerekes, Z. (2007). Public charity offer as a proximate factor of evolved reputation-building strategy: an experimental analysis of a real-life situation. *Evolution and Human Behavior*, 28, 277-284.
- Bolle, F. (2001). Why to buy your darling flowers: on cooperation and exploitation. *Theory and Decision*, 50, 1-28.
- Boone, J. L. (1998). The evolution of magnanimity: when is it better to give than to receive? *Evolution and Human Behavior*, 9, 1-21.
- Bower, J. (2005). The occurrence and function of victory displays within communication networks. In P. McGregor (Ed) *Animal Communication Networks*, pp. 114-126. Cambridge, UK: Cambridge University Press.
- Bshary, R. (2002). Biting cleaner fish use altruism to deceive image-scoring client reef fish. *Proceedings of the Royal Society of London B*, 269, 2087-2093.
- Bshary, R., & Grutter, A. (2006). Image scoring and cooperation in a cleaner fish mutualism. *Nature*, 441, 975-978.
- Burnham, T.C. & Hare, B. (2007). Engineering human cooperation: does involuntary neural activation increase public goods contributions? *Human Nature*, 18, 88-108.
- Buss, D. M., & Dedden, L. A. (1990). Derogation of competitors. *Journal of Social and Personal Relationships*, 7, 395-422.
- Coricelli, G., Joffily, M., Montmarquette, C., & Villeval, M. C. (2010). Cheating, emotions, and rationality: an experiment on tax evasion. *Experimental Economics*, 13, 226-247.
- Dabelsteen, T. (2005). Public, private or anonymous: facilitating and countering eavesdropping. In P. McGregor (Ed) *Animal Communication Networks*, pp. 38-62. Cambridge, UK: Cambridge University Press.
- Daly, M., & Wilson, M. (1988). *Homicide*. New York, NY: Aldine de Gruyter.

- DeClerck, C. H., Kiyonari, T., & Boone, C. (2009). Why do responders reject unequal offers in the Ultimatum Game? An experimental study on the role of perceiving interdependence. *Journal of Economic Psychology*, *30*, 35-343.
- Doutrelant, C., & McGregor, P. K. (2000). Eavesdropping and mate choice in female fighting fish. *Behaviour*, *137*, 1655-1669.
- Dunbar, R. I. M. (2006). Gossip in evolutionary perspective. *Review of General Psychology*, *8*, 100-110.
- Dunbar, R. I. M., Duncan, N. D. C., & Marriott, A. (1997). Human conversational behavior. *Human Nature*, *8*, 231-246.
- Ekström, M. (2011). Do watching eyes affect charitable giving? Evidence from a field experiment. *Experimental Economics*, *15*, 530-546.
- Ernest-Jones, M., Nettle, D. & Bateson, M. (2011). Effects of eye images on everyday cooperative behavior: a field experiment. *Evolution and Human Behavior*, *32*, 172-178.
- Feinberg, M., Willer, R., & Schultz, M. (2014). Gossip and ostracism promote cooperation in groups. *Psychological Science*, *25*, 656-664.
- Felson, R. B. (1978). Aggression as impression management. *Social Psychology*, *41*(3), 205-213.
- Fessler, D. M. T. (2002). Windfall and socially distributed willpower: the psychocultural dynamics of rotating savings and credit associations in a Bengkulu village. *Ethos*, *30*, 25-48.
- Fessler, D. M. T., Tiokhin, L. B., Holbrook, C., Gervais, M. M., & Snyder, J. K. (2014). Foundations of the Crazy Bastard Hypothesis: nonviolent physical risk-taking enhances conceptualized formidability. *Evolution and Human Behavior*, *35*, 26-33.
- Francey, D., & Bergmuller, R. (2012). Images of eyes enhance investments in a real-life public good. *PLOS ONE*, *7*, 1-7.
- Frank, R. H. (1988). *Passions Within Reason*. New York, NY: Norton.
- Frank, R. H. (2004). *What Price the Moral High Ground? Ethical Dilemmas in Competitive Environments*. Princeton, NJ: Princeton University Press.
- Getty, T. (2006). Sexually selected signals are not similar to sports handicaps. *Trends in Ecology and Evolution*, *21*, 83-88.
- Gintis, H., Smith, E. A., & Bowles, S. (2001). Cooperation and costly signaling. *Journal of Theoretical Biology*, *213*, 103-119.

- Goldstein, N. J., Griskevicius, V., & Cialdini, R. B. (2007). Invoking social norms: a social psychology perspective on improving hotels' linen-reuse programs. *Cornell Hospitality Quarterly*, 48, 145-150.
- Grafen, A. (1990). Biological signals as handicaps. *Journal of Theoretical Biology*, 144, 517-546.
- Griskevicius, V., Cantu, S., & Van Vugt, M. (2012). The evolutionary bases for sustainable behavior. *Journal of Public Policy and Marketing*, 31, 115-128.
- Griskevicius, V., Tybur, J.M., Gangestad, S.W., Perea, E. F., Shapiro, J. R., & Kenrick, D. T. (2009). Aggress to impress: hostility as an evolved context-dependent strategy. *Journal of Personality and Social Psychology*, 96(5), 980-994.
- Griskevicius, V., Tybur, J. M., Sundie, J. M., Cialdini, R. B., Miller, G. F., & Kenrick, D. T. (2007). Blatant benevolence and conspicuous consumption: when romantic motives elicit strategic costly signals. *Journal of Personality and Social Psychology*, 93(1), 85-102.
- Griskevicius, V., Tybur, J. M., & Van den Bergh, B. (2010). Going green to be seen: status, reputation, and conspicuous conservation. *Journal of Personality and Social Psychology*, 98(3), 392-404.
- Gurven, M., Allen-Arave, W., Hill, K., & Hurtado, A. M. (2000). "It's a Wonderful Life": signaling generosity among the Ache of Paraguay. *Evolution and Human Behaviour*, 21, 263-282.
- Haley, K.J. & Fessler, D.M.T. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behaviour*, 26, 245-256.
- Harbaugh, W.T. (1998). What do donations buy? A model of philanthropy based on prestige and warm glow. *Journal of Public Economics*, 67, 269-284.
- Hardy, C., & Van Vugt, M. (2006). Nice guys finish first: The competitive altruism hypothesis. *Personality and Social Psychology Bulletin*, 32, 1402-1413.
- Hawkes, K., & Bliege Bird, R. (2002). Showing off, handicap signaling, and the evolution of men's work. *Evolutionary Anthropology*, 11, 58-67.
- Hess, N. H., Hagen, E. H. (2006). Psychological adaptations for assessing gossip veracity. *Human Nature*, 17(3), 337-354.
- Hill, K., & Kaplan, H. (1988). Tradeoffs in male and female reproductive strategies among the Ache: part 1. In *Human Reproductive Behaviour: A Darwinian Perspective* (Eds. L Betzig, M. Borgerhoff Mulder, & P. Turke), 277-289. Cambridge, UK: Cambridge University Press.

- Hoffman, E., McCabe, K., Schachat, K., & Smith, V. (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior*, 7, 346-380.
- Johnstone, R. A., & Bshary, R. (2004). Evolution of spite through indirect reciprocity. *Proceedings of the Royal Society of London Series B*, 271, 1917-1922.
- Kazim, A. J. N., & Aureli, F. (2005). Redirection of aggression: multiparty signalling within a network? In P. McGregor (Ed) *Animal Communication Networks*, pp. 191-218. Cambridge, UK: Cambridge University Press.
- Keser, C. (2003). Experimental games for the design of reputation management systems. *IBM Systems Journal*, 42, 498-506.
- Kniffin, K. M., & Wilson, D. S. (2005). Utilities of gossip across organizational levels: multilevel selection, free-riders, and teams. *Human Nature*, 16, 278-292.
- Kraznow, M. M., Cosmides, L., Pedersen, E. J., & Tooby, J. (2012). What are punishment and reputation for? *PLOS ONE*, 7(9), e45662.
- Kurzban, R., & Houser, D. (2005). Experiments investigating cooperative types in humans: a complement to evolutionary theory and simulations. *PNAS*, 102, 1803-1807.
- Lyle, H. F. III, Smith, E. A., & Sullivan, R. J. (2009). Blood donations as costly signals of donor quality. *Journal of Evolutionary Psychology*, 4, 263-286.
- Matsumoto, D., & Hwang, H. S. (2012). Evidence for a nonverbal expression of triumph. *Evolution and Human Behavior*, 33, 520-529.
- Matos, R. J., & Schlupp, I. (2005). Performing in front of an audience: signallers and the social environment. In P. McGregor (Ed) *Animal Communication Networks*, pp. 63-83. Cambridge, UK: Cambridge University Press.
- McCullough, M. E., Kurzban, R., & Tabak, B. A. (2013). Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences*, 36, 1-58.
- McGregor, P. K., & Peake, T. M. (2000). Communication networks: social environments for receiving and signalling behaviour. *Acta ethologica*, 2, 71-81.
- McNamara, J. M., Barta, Z., Frohmage, L., & Houston, A. I. (2008). The coevolution of choosiness and cooperation. *Nature*, 451, 189-192.
- McNamara, J. M., & Houston, A. I. (2002). Credible threats and promises. *Philosophical Transactions of the Royal Society of London: Series B*, 357, 1607-1616.
- Mifune, N., Hashimoto, H. & Yamagishi, T. (2010). Altruism toward in-group members as a reputation mechanism. *Evolution and Human Behavior*, 31, 109-117.

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

Milinski, M., Semmann, D., & Krambeck, H.-J. (2002). Reputation helps solve the “tragedy of the commons”. *Nature*, 415, 424-426.

Milinski, M., Semmann, D., Krambeck, H.-J., & Marotzke, J. (2006). Stabilizing the Earth’s climate is not a losing game: supporting evidence from public goods experiments. *PNAS*, 103, 394-3998.

Nesse, R. M. (2007). Runaway social selection for displays of partner value and altruism. *Biological Theory*, 2(2), 143-155.

Nettle, D., Nott, K., & Bateson, M. (2012). “Cycle thieves, we are watching you”: impact of a simple signage intervention on bicycle theft. *PLOS ONE*, 7, e51738.

Noë, R., & Hammerstein, P. (1994). Biological markets: supply and demand determine the effect of partner choice in cooperation, mutualism and mating. *Behavioral Ecology & Sociobiology*, 35, 1-11.

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

Nowak, M. A., Page, K. M., & Sigmund, K. (2000). Fairness versus reason in the Ultimatum Game. *Science*, 289, 1773-1775.

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

Ohtsubo, Y., & Watanabe, E. (2009). Do sincere apologies need to be costly? Test of a costly signaling model of apology. *Evolution and Human Behavior*, 30, 114-123.

Ohtsuki, H., & Iwasa, Y. (2004). How should we define goodness? Reputation dynamics in indirect reciprocity. *Journal of Theoretical Biology*, 231, 107-120.

Ohtsuki, H., & Iwasa, Y. (2006). The leading eight: Social norms that can maintain cooperation by indirect reciprocity. *Journal of Theoretical Biology*, 239, 435-444.

Ohtsuki, H., & Iwasa, Y. (2007). Global analyses of evolutionary dynamics and exhaustive search for social norms that maintain cooperation by reputation. *Journal of Theoretical Biology*, 244, 518-531.

Oliveira, R. F., McGregor, P. K., & Latruffe, C. (1998). Know thine enemy: fighting fish gather information from observing conspecific interactions. *Proceedings of the Royal Society of London B*, 265, 1045-1049.

Ophir, A. G., Persaud, K. N., & Galef, B. G. (2005). Avoidance of relatively aggressive male Japanese quail (*Coturnix japonica*) by sexually experienced conspecific males. *Journal of Comparative Psychology*, 119, 3-7.

Otter, K. A., McGregor, P. K., Terry, A. M. R., Burford, F. R. L., Peake, T. M., & Dabelsteen, T. (1999). Do female great tits (*Parus major*) assess males by eavesdropping? A field study using interactive song playback. *Proceedings of the Royal Society of London Series B*, 266, 1305-1309.

Otter, K. A., Stewart, I. R. K., McGregor, P. K., Terry, A. M. R., Dabelsteen, T., & Burke, T. (2001). Extra-pair paternity among Great tits *Parus major* following manipulation of male signals. *Journal of Avian Biology*, 32, 338-344.

Panchanathan, K., & Boyd, R. (2004). Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature*, 432, 499-502.

Peake, T. M. (2005). Eavesdropping in communication networks. In P. McGregor (Ed.) *Animal Communication Networks*, pp. 13-37. Cambridge, UK: Cambridge University Press.

Pfeiffer, T., Tran, L., Krumme, C., & Rand, D. G. (2012). The value of reputation. *Journal of the Royal Society Interface*, 9, 2791-2797.

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

Pinker, S. (2011). *The Better Angels of Our Nature: Why Violence Has Declined*. New York, NY: Viking.

Pizarro, D. A., Tannenbaum, D., & Uhlmann, E. (2012). Mindless, harmless, and blameworthy. *Psychological Inquiry*, 23, 185-188.

Price, M. E. (2003). Pro-community altruism and social status in a Shuar village. *Human Nature*, 14, 191-208.

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

Roberts, G. (1998). Competitive altruism: from reciprocity to the handicap principle. *Proceedings: Biological Sciences*, 265, 427-431.

Roberts, G. (2008). Evolution of direct and indirect reciprocity. *Proceedings: Biological Sciences*, 275, 173-179.

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

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

Sell A, Tooby J, Cosmides L. 2009. Formidability and the logic of human anger. *PNAS*. 106(35):15073-15078.

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

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

Sommerfeld, R. D., Krambeck, H.-J., Semmann, D., & Milinski, M. (2007). Gossip as a alternative for direct observation in games of indirect reciprocity. *PNAS*, 104, 17435-17440.

Smith, E. A. (2004). Why do good hunters have higher reproductive success? *Human Nature*, 15, 343-364.

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

Smith, E. A., & Bliege Bird, R. (2000). Turtle hunting and tombstone opening: public generosity as costly signaling. *Evolution and Human Behavior*, 21, 245-262.

Smith, E. A., & Bliege Bird, R. (2005). Costly signalling and cooperative behaviour. In *Moral Sentiment and Material Interests: The Foundations of Cooperation in Economic Life* (Eds. H. Gintis, S. Bowles, R. Boyd, & E. Fehr) pp. 115-148. Cambridge, MA: MIT Press.

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

Sommerfeld, R. D., Krambeck, H.-J., & Milinski, M. (2008). Multiple gossip statements and their effect on reputation and trustworthiness. *Proceedings of the Royal Society of London Series B*, 275, 2529-2536.

Sosis, R. (2004). The adaptive value of religious ritual. *American Scientist*, 92, 166-172.

Sparks, A., & Barclay, P. (2013). Eyes increase generosity, but not for long: the limited effect of a false cue. *Evolution & Human Behavior*, 34, 317-322.

Sparks, A., & Barclay, P. (in preparation). Condemnation alone does not sustain long-term cooperation, but tangible punishment does.

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

Tannenbaum, D., Uhlmann, E L., & Diermeier, D. (2011). Moral signals, public outrage, and immaterial harms. *Journal of Experimental Social Psychology*, 47, 1249-1254.

Van Vugt, M., Roberts, G., & Hardy, C. (2007). Competitive altruism: a theory of reputation-based cooperation in groups. In R Dunbar & L Barrett (Eds.) *Oxford Handbook of Evolutionary Psychology*, pp. 531-540. Oxford, UK: Oxford University Press.

Veblen, T. (1899/1994). *The Theory of the Leisure Class*. Penguin Classics.

Wedekind, C., & Braithwaite, V. A. (2002). The long-term benefits of human generosity in indirect reciprocity. *Current Biology*, 12, 1012-1015.

West-Eberhard, M. J. (1979). Sexual selection, social competition, and evolution. *Proceedings of the American Philosophical Society*, 123, 222-234.

West-Eberhard, M. J. (1983). Sexual selection, social competition, and speciation. *The Quarterly Review of Biology*, 58, 155-183.

Willer, R. (2009). Groups reward individual sacrifice: the status solution to the collective action problem. *American Sociological Review*, 74, 23-43.

Yamagishi, T., Horita, Y., Takagishi, H., Shinada, M., Tanida, S., & Cook, K. S. (2009). The private rejection of unfair offers and emotional commitment. *PNAS*, 106(28), 11520-1523.

Zahavi, A. (1977). Reliability in communication systems and the evolution of altruism. In *Evolutionary Ecology* (Eds. B Stonehouse and C. Perrins), pp. 253-259. Baltimore, MD: University Park Press.