

Reputation management in the age of the world-wide web

Claudio Tennie¹, Uta Frith^{2,3} and Chris D. Frith^{2,4}

¹ Max Planck Institute for Evolutionary Anthropology, Department of Developmental and Comparative Psychology, Deutscher Platz 6, D-04103 Leipzig, Germany

² Interacting Minds Project, Center for Functionally Integrative Neuroscience, Aarhus University Hospital, 800 Aarhus C, Denmark

³ Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, UK

⁴ Wellcome Trust Centre for Neuroimaging at University College London, 12 Queen Square, London WC1N 3BG, UK

The reciprocal interactions with others that play such a significant part in our lives depend upon trust; individuals need to be confident that their partners are cooperative, and that they will return favours. Reputation permits the choice of better partners and provides incentives to be more cooperative. These uses of reputation are not unique to humans. However, in complex human societies, with large numbers of potential partners, keeping track of each other's reputation is a vital part of everyday life, and, in an inevitable arms race, ever more powerful strategies of reputation management are being developed. In this article, we bring together insights from different disciplines to throw new light onto the importance and scope of reputation management.

The importance of reputation

In the age of Google, Facebook and eBay, you are likely to interact on a daily basis with many people about whom you know nothing. By contrast, in small communities you would have first-hand knowledge about the reputation of everyone with whom you interacted. What enables individuals to track reputation even in complex societies, and why is it so important? In the Age of Enlightenment, Adam Smith [1] came to the surprising conclusion that an individual is born to value a good reputation even more than profit. In this paper, we review recent evidence from comparative biology, behavioural economics and social cognitive neuroscience to explain why this might be so. We do this with reference to the evolutionary roots and the still evolving strategies for managing and monitoring reputation in complex societies (see Boxes 1–3).

What is reputation and what is it for?

In evolutionary theories, reputation is one of many signals providing information about the likely behaviour of an individual (see section 7.5 in reference [2]). Such signals can apply to different aspects of behaviour, such as skill or dominance; however, the focus in this review is on behaviour in reciprocal interactions, and on reputation as a signal of the cooperativeness and reciprocity that are such important features of human interactions. In this context, reputation has a key role in maintaining social cohesion. It enables incentive mechanisms for reducing uncooperative

behaviour (free riding) [3], and guides the choice of partners in a social network. Studying the management of reputation is of particular relevance in contemporary

Glossary

Altruism: Performing a beneficial act for another individual that is costly to the direct fitness of the performer. If behaviour acts as an investment leading to an individual's long-term gain, it is selfish rather than altruistic.

Altruistic punishment: Action taken by an individual to punish a free rider at a cost to the punisher. This leads to a gain for the group, but might be altruistic at the level of the individual [13]. Alternatively, punishers might gain directly by receiving reputation-mediated benefits for their costly actions.

Audience effect: An individual's behaviour changes in the presence of an observer (usually a member of the same species, conspecific). This effect has been documented in a large variety of species. In human social psychology, the audience effect was first described by Zajonc [56], who proposed that it produced an increase in arousal.

Image scoring: Evaluating the reputation of individuals in a group. This strategy is crucial in indirect reciprocity, because partners have to discriminate between those who cooperate and those who do not [7].

Mutual benefit: Occurs when two individuals benefit immediately from a cooperative interaction.; for example, a simultaneous trade in which each individual knows that the object he is getting is of more benefit to him than the object he is giving. Mutual benefit can also occur across species (mutualism), as with the gut flora that aid human digestion.

Byproduct mutualism: For example, chimpanzees might join group hunts because it directly increases their chances of gathering enough meat to satisfy their own nutritional needs [57]. By joining a hunt, an animal might also be close enough to harass the successful hunter [58]. Such behaviour need not, therefore, be an example of reciprocity.

Reciprocity: Individual A performs a beneficial act (costly to A) for individual B, in order to increase the probability of receiving direct future benefits. The investor recoups her investment either from her original interaction partner (A helps B; subsequently B helps A: direct reciprocity [59]) or through a beneficial interaction from some other individual (A helps B; subsequently C helps A: indirect reciprocity [60]). For such behaviour to become more frequent during evolution or the life time of the individual, the short-term costs of investing must be exceeded by the long-term benefits.

Pseudoreciprocity: Individual A performs a beneficial act for individual B, in order to increase the probability of receiving incidental benefits from B's future actions. Because the return benefits to A derive from behaviours that B performs to benefit B, not A, there is no motivation for cheating in pseudoreciprocity.

Manipulation/exploitation: This behaviour can often explain apparent mutualism. For example, chimpanzees might 'share' meat with each other to avoid harassment costs produced by beggars [8,61]. Similar pre-emptive avoidance might explain why chimpanzees sometimes 'help' conspecifics or humans [62,63], and fail to force others to work for them [64].

Corresponding author: Frith, C.D. (c.frith@ucl.ac.uk).

Box 1. Reputation management in fish

The reef fish *Labroides dimidiatus* interacts as a cleaner with other, larger reef fish, which is mutually beneficial. However, the cleaner fish has the option of feeding on the mucus of the client fish rather than eating the parasites. Because the cleaners prefer to eat the mucus [65], resisting this temptation and restricting their feeding to the parasites is a form of investment. In the long run, this behaviour pays off because cleaners that do not bite are chosen by more clients [66].

When a client fish approaches a cleaner, it trusts that cheating will not occur. If the cleaner does cheat by biting the client, then the client makes a short jolt and terminates the interaction by shaking off and possibly chasing away the cleaner (introducing the incentive to avoid punishment). By observing this behaviour, other potential clients can select which cleaner fish to approach. Clients are most likely to approach cleaners that have just had an interaction that terminated well. They are less likely to approach cleaners that they have not observed interacting, and they are least likely to approach cleaners that have just had an interaction that terminated badly [66]. A cleaner that behaves altruistically with one client is more likely to be approached by another client. This is an example of indirect reciprocity, as cheating can occur and provides the preconditions for reputation management.

One example of such reputation management is when cheating cleaner fish generate signals that they are altruistic (tactical deception). Cleaner fish will behave altruistically (not biting) with a first client to get access to the second, 'observer', client, on whom they may then cheat [67]. Another example involves cleaner fish that sometimes stimulate the backs of their clients with their fins [68]. Because their mouths point away from the client, this behaviour is incompatible with feeding. Bshary [66] suggests that this behaviour acts as a signal to observing clients, indicating that this is a cooperative cleaner. Consistent with the idea that this frequently constitutes tactical deception, he observed that, after biting cleaners had used tactile stimulation on one client, they were more likely to bite their next client.

Box 2. Is there reputation management in chimpanzees?

In contrast to the case of cleaner fish, the evidence from the only two studies on the use of reputation and image scoring in chimpanzees is inconclusive. The first study found that, after third party demonstrations of either food-providing or non-food-providing humans, chimpanzees, but not other great apes, choose the proximity of the food-providing humans [69]. However, food providing was not contrasted with a neutral demonstration, but instead with demonstrations of dominance (non-food-providers slapped the hands of the third party). Thus, the effect could have been driven by the chimpanzees using the dominance information alone (namely to avoid these demonstrators). The second study used a very similar setup, although without the confound of dominance information [70]. However, in this situation, the chimpanzees did not differentiate between the demonstrators. Later experiments in that study did find differential effects, but subjects had then been exposed to what is best described as heavy training regimes. In sum, chimpanzees use reputation signals about dominance [71], but they do not seem to use reputation signals about reciprocal cooperativeness.

However, there is also very little evidence of reciprocity in chimpanzees. As reciprocity with the possibility of cheating is a precondition for reputation management, we might expect that reputation management would be of little interest to this species. Chimpanzees show some cooperative interactions, but both field and captive data suggest that these are mostly non-reciprocal kinds of cooperation, which are either driven by kin selection, or which can be explained by mutualism and pseudoreciprocity (see Glossary). Two controlled tests for reciprocity in chimpanzees yielded no evidence [62,72], and the only weakly positive study still awaits replication [64]. It is therefore difficult to find examples of behaviour in which reputation and image scoring would have a role [4,8].

Box 3. Outstanding questions

How can mechanisms underlying reputation be modelled in large communities, in which contact is fleeting and indirect? With increasing social interactions on the internet, there is a pressing need to build workable systems for evaluating trust and discovering fraud as rapidly as possible. Could there be a reputation-tracking mechanism that goes beyond summing up positive and negative feedback evaluations of individuals provided by those who have first-hand interactions with them? How can the reputation of those who provide this feedback be tracked?

How can reputation held by institutions rather than individuals be modelled? Can the management of reputation via institutions be achieved by treating the institution as if it were an individual? What signals for reputation are used by institutions, and do they differ from signals used by individuals?

To what extent does reputation management depend upon mentalizing? Do individuals with autism care about their own reputation [73]? It is not yet known as to what extent they are subject to the audience effect. If individuals with autism spectrum disorder have a mentalizing deficit, their strategies for managing reputation and their evaluation of reputation in others would be particularly informative about the role of mentalizing in reputation management.

Reputation in other species? Many species show an audience effect (e.g. the vocalisations of some birds change in the presence of other birds [74,75]). Are these species more likely than others to develop reputation management strategies? Are mechanisms underlying reputation management used, for instance, by cleaner fish and client fish in the same manner as in humans?

Reputation in young children? Audience effects on smiling behaviour have been shown in 10-month-olds [76], but when do young children first become concerned about the opinion of others, and how do they learn strategies of reputation management?

The salience of reputation. Have reputation signals become more important than the behaviour that these signals are supposed to indicate? Our impression is that reputation management has become an industry in its own right, and might risk losing touch with the purpose of gaining a good reputation. This would be the case if firms invest more in reputation management than in the quality of their goods or services. What consequences does this have for the reputation arms race?

human societies, where the large size of the interacting groups increases anonymity, reducing the ability to monitor reputation.

Given the focus of this review, it is important to state what is meant by reciprocity. The key feature of reciprocity (not to be confused with mutual benefit [4]) is that the benefit for one of the individuals involved is not immediate. In a reciprocal interaction, an individual does something that is beneficial for another individual, while incurring a cost to the performer. Such behaviour is a form of altruism (at least in the short-term; see Glossary), because it is disadvantageous to the investor. The challenge for evolutionary theories is to explain how such cooperative behaviour can evolve. Different theories have provided different answers to this question. Competitive altruism theory [5,6] and indirect reciprocity theory [7] propose that the cooperative individual benefits in the long term through increased cooperation from others. In this framework, the costly initial behaviour is a form of investment [8], because the individual expects something in return in the future (see Glossary). Strong reciprocity theory [9], in contrast, proposes that human cooperation evolved because groups containing many cooperative individuals have advantages over other groups. Within all these approaches, however, are situations in which reputation

is needed to establish cooperation. Uncooperative behaviour results in a bad reputation, which is associated with several disadvantages. As a result, reputation is a stabilising factor, reducing the likelihood of uncooperative behaviour.

Mechanisms of reputation and reputation management

In social animals exhibiting reciprocal behaviour, it is likely that there are at least two mechanisms of reputation and reputation management, one providing a basis for selecting appropriate partners (choice) and one providing a basis for rewarding them (incentives). Thus, individuals who have gained a reputation for cooperative behaviour are more likely to be chosen as partners [10], whereas those with a poor reputation are likely to be shunned and excluded [11]. Similarly, individuals who have gained a reputation for cooperative behaviour are likely to receive rewards [12], whereas those with a poor reputation might be punished [13]. These mechanisms provide strong motivation to be cooperative, even for individuals with a natural tendency to be uncooperative [14] (see [Box 1](#) for an example in fish).

Reputation counts for little if it is not signalled. It is therefore the signals of a good reputation that guide choosing with whom to cooperate and when. However, because signals are not the same as the behaviour they signal, discrepancies can occur. These can be due to unintentional errors, but also to exaggerated displays and outright deception. Thus, mechanisms are also needed for monitoring the actions of individuals, which detect discrepancies between reputation and behaviour.

In humans, one such mechanism is provided by the ability to mentalize. Mentalizing refers to the pervasive ability to attribute intentions to others and to predict their behaviour on the basis of their mental states [15]. The neural substrate of this ability is active during tasks that require the judgement of a person's intentions, and during performance of iterative versions of economic games, such as the 'ultimatum game' and the 'prisoners' dilemma' [16]. Through simultaneously scanning two partners in a trust game, a first glimpse has been gained of the neural mechanisms underlying reciprocity and trust in humans [17]. A unique response pattern in the caudate nucleus was identified as a signal of the intention to trust. This signal became anticipatory when players could make accurate guesses about their partner's investment return, and is thus just what one would expect from a dynamic representation of a partner's reputation.

A further mechanism relevant for the monitoring of reputation in humans is epistemic vigilance [18]. This automatic tendency to care about and if necessary query the validity of information obtained from other people, is present even in young children. Thus, when learning something new, 3–4-year olds spontaneously prefer to learn from a previously reliable informant rather than a previously unreliable person. [19,20].

Direct and indirect sources of reputation

Direct observations of cooperative behaviour and/or of the signals of such behaviour are a major source for determining reputation. The 'audience effect' suggests that the

individual being judged automatically takes into account observers who are potential partners. Social animals, who invest in others, send out suitable signals when observed, but will not do so when unobserved [21]. These signals incur a cost, and the more costly they are, the more they should enhance reputation [22]. Costly reputation-enhancing behaviour tends to occur when many observers are present who not only can take notice of the behaviour, but can also spread the news. Likewise, reputation-diminishing behaviour, for example cheating or exploitation of others' investments, tends to be performed when no audience is present [23,24].

The audience effect and effects of anonymity are two sides of the same coin, working in opposite directions. When there is anonymity, and this is often the case with large groups, it is hard to track individual reputation, and free riders can invade more easily [21]. Removing anonymity and reinstating an audience will allow reputation to be acquired again, and will lead to increases in cooperation [13,25].

Indirect sources of reputation, for instance provided by gossip (reported reputation [26]), are particularly important in human societies. Gossip requires the ability to attribute knowledge and ignorance to other people, and to monitor what they know and what they don't know. Hence, this tool of reputation management strongly depends on our ability to mentalize [27]. Humans spontaneously track which information is new to particular recipients and selectively impart this knowledge [28]. Not all knowledge can be trusted equally, and continuous checking of the reliability of the source of new information is advantageous. Here, epistemic vigilance (can I trust the new information imparted to me?) can be co-opted for the purpose of monitoring the reputation of others (can I trust this potential partner?). Given a tool for the ready evaluation of 'hearsay' information, the dissemination of reputation in large societies can be very rapid.

The importance of indirect information about the reputation of others over and above direct observation was also confirmed in a neuroimaging study [29]. This study used vignettes to define the reputation of partners before volunteers engaged in an economic game, in which these partners did or did not behave according to the image presented in the vignette. The effects of this hearsay evidence, observed in behaviour and brain activity, suggested that more weight was placed on the reported reputation than on the actual behaviour of the partner. Gossip often pools information from many sources and thus should, other things being equal, be preferred over restricted personal information. Advantages of relying on pooled sources of information were recently demonstrated in a social learning strategies tournament [30].

Managing modern risks of reputation management

As information flows ever more rapidly and efficiently as a result of the worldwide web, we can observe evolving strategies of reputation management. There are ever larger groups within which to interact, resulting in more partner choices but also a greater degree of anonymity. With this increased choice, there are also increased opportunities for deceptive use of reputation, threatening the

maintenance of cooperative behaviour. Trading via the internet is becoming increasingly common. When purchasing or bidding for products, these products cannot be fully evaluated before purchase [31], and there is a risk that the seller is dishonest and will substitute an inferior product or never send it at all [32]. To monitor the honesty of transactions, the trader's reputation needs to be monitored. For example, the online bidding company eBay has introduced an online rating system explicitly stating: 'Feedback is each user's reputation on eBay. It fosters trust between people by acting as both an incentive to do the right thing and as a mark of distinction for those who conduct transactions with respect, honesty and fairness.' On this basis, bidders can rate the sellers after trading, and these ratings are constantly updated for everyone to see. Buyers can then choose sellers with high ratings, and there is an incentive for sellers to acquire high ratings. In theory then, if reputation management works well, a dishonest trader's reputation will be ruined; this news will spread over the market place, and his business will suffer. However, traders can retaliate. For example, they can pose as buyers who deliberately give poor ratings to diminish the reputation of other sellers, which might have the effect of reducing competition (see Figure 1).

Another threat to reputation monitoring is the ease of setting up new identities. If an individual acquires a poor reputation, they can cease trading from that address and start again with a different identity, thus circumventing the problem of having to live with – or having to improve – a bad reputation. Individuals can also create more than one identity, and then use one of these identities to enhance the reputation of the other, thus subverting the control of reputation by report (so-called 'sock-puppeting') [33]. Such individuals are effectively freed from reputational consequences [32,34]. It should be noted that, realizing this, some internet companies (such as eBay) have installed systems to prevent users from establishing new identities or to mark them as having recently changed identity.

Costs in the arms race

One indicator of an honest signal of reputation is its cost. A costly investment is a signal that is difficult to fake. The use of costly signals as trustworthy guides to reputation is revealed clearly in organised crime, because criminals cannot advertise their skills through the normal channels [35]. Furthermore, criminal groups are aware that their organisations are targets for infiltration by undercover agents. The ultimate costly signal for a criminal wishing to be accepted by the mafia is the requirement for him to kill someone before he can become 'a made guy'. Although this requirement holds other benefits for the mafia boss, such as locking the individual into the family, it also acts as a crucial test; few government agents would commit murder to gain a trustworthy reputation among criminals [36].

A problem arises if there is a price dump in signals that used to be costly [37]. Many signals can now be produced inexpensively by digital means. As a result, signals that used to be honest can no longer be trusted. An example is the ease of airbrushing photographs. This increasing ease of manipulating the media has led to the untrustworthy control of reputation through 'spin-doctoring' [38].

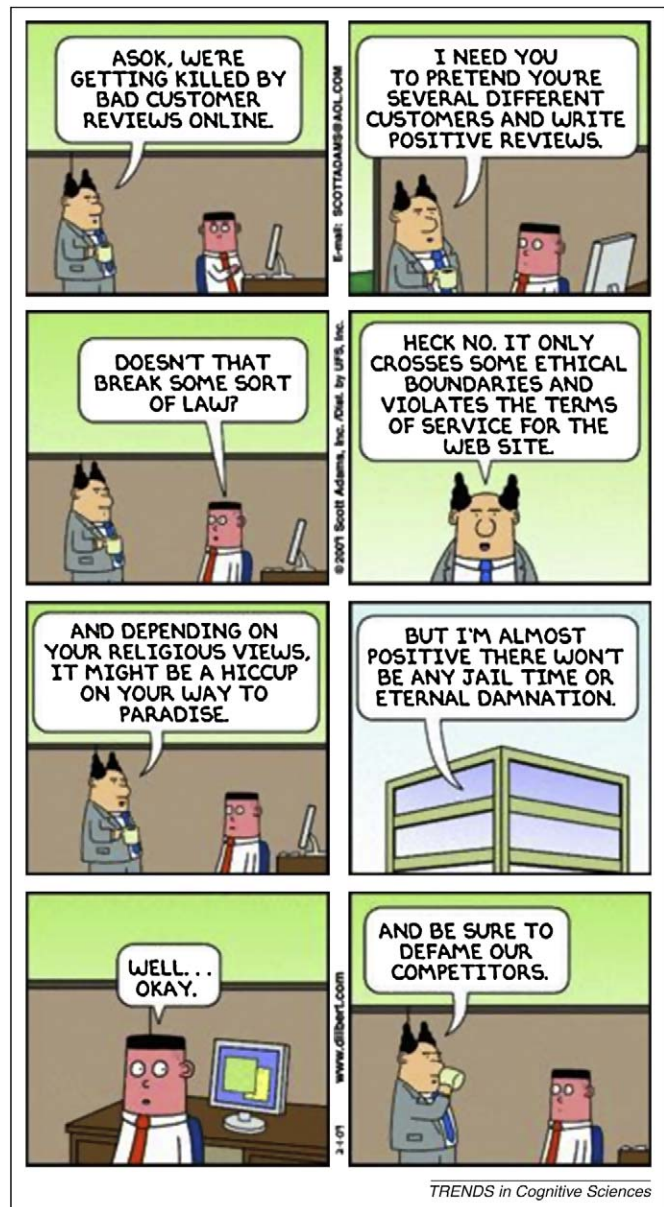


Figure 1. The comic strip 'Dilbert' (©Scott Adams), illustrating the unethical use of reputation management on the internet.

There is also a mental cost in keeping track of the reputation of many individuals [39,40]. The problem of tracking multiple reputations increases the need for automatic evaluation of additional knowledge from third parties. In some situations, it seems not to matter to reputation judges whether reporters of reputation have a bad reputation for cooperation themselves [26]. Thus, it pays for those with a bad reputation to provide reliable information about others' reputation, as this might help them to rescue their own reputation.

Is direct face to face interaction always better than dealing with the anonymous multitude of individuals on Facebook and similar sites? Not necessarily. When direct observation is available, humans watch for cues of dishonesty. Such cues are found in non-verbal behaviour [41] and even in the shape (male) of the face [42], but these cues can be subverted. Trivers [43] suggested that deception might be more successful if reputation holders deceive them-

selves about their reputation, as this would allow non-verbal behaviour to reflect honestly – if hypocritically – held beliefs. This kind of self-deception can be seen in the ‘illusory superiority’ displayed by many people, considering themselves to be better than average in their driving ability, sense of humour and many other desirable attributes [44]. Another self-serving bias is the tendency to rate a transgression committed by oneself as less objectionable than the same transgression committed by another [45]. Such biases will create the illusion that one’s own reputation is better than average and therefore might be able to withstand censure by others when one behaves in an uncooperative manner.

Evolving strategies: the role of institutions

In complex societies, there is the danger that reputation takes on a life of its own and diverges more and more from actual behaviour. This is not only the result of deceptive intentions, but can also occur unintentionally. With large numbers of potential partners, divergence might simply be due to errors of memory or identity. How can such errors be minimised? One strategy is to use multiple independent sources for checking information about someone’s reputation [46]. From here, it is a small step towards representing reputation in a distributed fashion, not solely in individuals, but in institutions [47].

This strategy long predates the appearance of the internet. Guilds and professional societies are obvious examples. We might know nothing of doctors who treat us, but we trust them because they are on the medical register. In this case it is the register that has reputation, but, by being on the register, an individual acquires the same reputation. Here again, people will use the register to choose their doctor, and thus there is a strong incentive for doctors to behave in such a way as to remain on the register. In the long run, the reputation will be determined by the behaviour of all the individuals on the register. To guard the reputation of such a group, it is necessary to have regulation through incentives and choice. For example, Gürerk and colleagues [48] demonstrated experimentally a clear competitive advantage of and preference for sanctioning over non-sanctioning institutions.

However, new problems arise when reputation belongs to an organisation rather than to an individual. In this case, the wider community needs to retain the power to provide incentives and choice, and the ability to disband the organisation if it no longer applies sanctions to its members. The case of the recent scandal of child abuse by priests who are perceived as not sufficiently punished shows that even the reputation of an institution as powerful and ancient as the Catholic Church can be shaken. In such cases, strong attempts are made to attach blame to specific individuals. If no specific individuals can be practically linked with the organisation or made to take responsibility (separation of entity from identity [34]), the application of punishment and exclusion is difficult, and the reputation of the organisation remains uncertain.

A major problem arises when an organisation with a previously good reputation becomes predatory. For instance, as retirement approaches, the owner of an organisation might be willing to lose his reputation in order to

obtain a large short-term gain. Such behaviour is kept in check through the possibility of selling on the organisation, which will be more valuable if it has a high reputation [34]. However, the possibility that an organisation can be sold raises new risks. A company’s good name might have been deliberately acquired by a group of unscrupulous individuals who have made this sort of takeover their speciality, to turn past reputation into quick present profit [49]. Through their specialisation, predators find it inefficient to build a good reputation on their own. Thus, they might have strong incentives to outbid cooperative individuals in the market for good names [34]. In this case, the reputation system breaks down.

The future of reputation management

Does our ever-widening social network and the associated increase in anonymity inevitably increase selfishness over reciprocity? Using computer simulations of cooperative games, Suzuki and Akiyama [50] confirmed that the frequency of cooperation can decrease when group size increases. They note, however, that cooperation can coexist in stable equilibrium with betrayal, especially in medium sized groups.

Even though it is reliably reduced, cooperation need not break down completely in anonymous situations [50]. Individuals will often make costly investments at the outset of experiments on cooperation even under conditions of anonymity, which is not to be expected from an agent driven by rational self-interest [51]). These effects might reflect strong reciprocity [9], but could also occur as the result of a mismatch between the circumstances in which cooperative behaviour evolved and the form taken by the experiment [52]. For example, however explicit the instructions for these experiments, people cannot discount the possibility that their behaviour might be observed and that the interaction might be repeated. Evidence for such apparently irrational behaviour is seen in the experiment of Bateson *et al.* [53], in which placing a photo of a pair of eyes above the ‘honesty box’ in the staff common room was sufficient to increase significantly the amount of money that was collected [54]. There are many real-life examples of investments being made without any guarantee of an audience, such as when anonymous donations are made to charity [55]. One might invoke an all-seeing God, or perhaps the need to build a reputation in one’s own eyes, which is captured by the term ‘self-respect’. If there continues to be a perceived risk of an audience being present, then this alone would explain why the baseline is for people to cooperate rather than not to cooperate. This and other issues are now ripe for experimental investigation (Box 3).

Concluding remarks

There is a never-ending arms race between the desire to act selfishly and the need to achieve a good reputation. We have argued that mechanisms allowing choice of cooperative partner and sanctions against uncooperative partners have their roots in evolution and persist even today. We focus on the complexity of this arms race in the increasingly large social networks present in human societies, leading to ever-evolving strategies for reputation management. We propose that both this complexity and attempts

to master it derive from the ability to mentalize. Although incentive and choice still drive the desire for a good reputation, mentalizing drives the ability to detect the intentions behind actions. It also drives gossip, through which reputation can be created and destroyed, as well as the continuous vigilance through which gossip can be verified. The internet has dramatically increased the size of interacting groups, which are essentially anonymous. However, it also provides the tools for tracking vast numbers of reputations simultaneously. This can overcome the problem of excess cognitive load, and helps to reduce uncertainty by pooling input about reputation from multiple sources. Thus, modern human life provides ever more sophisticated means of managing reputation and an ever-greater need to understand the mechanisms of reputation management.

Acknowledgements

We are grateful to the participants in the Reputation Management Workshop held at the University of Aarhus in December 2009, supported by the AHRC CNCC scheme AH/E511112/1 and the Danish National research Council. For comments on earlier versions of this manuscript we thank Amrisha Vaish, Anke Bullinger, Kevin Langergraber, Emily Wyman and Guido Olschewski. We are also grateful to Christophe Heintz and the anonymous reviewers who generously helped us to improve this paper.

References

- Smith, A. *The Theory of Moral Sentiments* (Raphael D.D. and Macfie, A.L, eds), Liberty Classics
- Maynard Smith, J. and Harper, D. (2003) *Animal Signals*, Oxford University Press
- Fehr, E. and Fischbacher, U. (2004) Social norms and human cooperation. *Trends Cogn. Sci.* 8, 185–190
- West, S.A. *et al.* (2007) Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J. Evol. Biol.* 20, 415–432
- McNamara, J.M. *et al.* (2008) The coevolution of choosiness and cooperation. *Nature* 451, 189–192
- Roberts, G. (1998) Competitive altruism: from reciprocity to the handicap principle. *Proc. R. Soc. London B* 265, 427–431
- Nowak, M.A. and Sigmund, K. (1998) Evolution of indirect reciprocity by image scoring. *Nature* 393, 573–577
- Bshary, R. and Bergmuller, R. (2008) Distinguishing four fundamental approaches to the evolution of helping. *J. Evol. Biol.* 21, 405–420
- Fehr, E. *et al.* (2002) Strong reciprocity, human cooperation, and the enforcement of social norms. *Hum. Nat.* 13, 1–25
- Sylwester, K. and Roberts, G. (2010) Cooperators benefit through reputation-based partner choice in economic games. *Biol. Lett.* doi: 10.1098/rsbl.2010.0209 rsbl20100209
- Panchanathan, K. and Boyd, R. (2004) Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature* 432, 499–502
- Wedekind, C. and Milinski, M. (2000) Cooperation through image scoring in humans. *Science* 288, 850–852
- Fehr, E. and Gächter, S. (2002) Altruistic punishment in humans. *Nature* 415, 137–140
- Fehr, E. *et al.* (2009) A Behavioral account of the labor market: the role of fairness concerns. *Annu. Rev. Econ.* 1, 355–384
- Frith, U. and Frith, C.D. (2003) Development and neurophysiology of mentalizing. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 358, 459–473
- Rilling, J.K. *et al.* (2004) The neural correlates of theory of mind within interpersonal interactions. *Neuroimage* 22, 1694–1703
- King-Casas, B. *et al.* (2005) Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83
- Mascaro, O. and Sperber, D. (2009) The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition* 112, 367–380
- Birch, S.A. *et al.* (2008) Three- and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition* 107, 1018–1034
- Nurmsoo, E. and Robinson, E.J. (2009) Identifying unreliable informants: do children excuse past inaccuracy? *Dev. Sci.* 12, 41–47
- Andreoni, J. and Bernheim, B.D. (2009) Social image and the 50-50 norm: a theoretical and experimental analysis of audience effects. *Econometrica* 77, 1607–1636
- Semmann, D. *et al.* (2004) Strategic investment in reputation. *Behav. Ecol. Sociobiol.* 56, 248–252
- Rege, M. and Telle, K. (2004) The impact of social approval and framing on cooperation in public good situations. *J. Public Econ.* 88, 1625–1644
- Soetevent, A.R. (2005) Anonymity in giving in a natural context—a field experiment in 30 churches. *J. Public Econ.* 89, 2301–2323
- Milinski, M. *et al.* (2002) Reputation helps solve the 'tragedy of the commons'. *Nature* 415, 424–426
- Sommerfeld, R.D. *et al.* (2007) Gossip as an alternative for direct observation in games of indirect reciprocity. *Proc. Natl. Acad. Sci. U. S. A.* 104, 17435–17440
- Frith, U. and Frith, C. (2010) The social brain: allowing humans to boldly go where no other species has been. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 365, 165–176
- Saylor, M.M. *et al.* (2006) Telling others what's new: preschoolers' adherence to the given-new contract. *J. Cogn. Dev.* 7, 341–379
- Delgado, M.R. *et al.* (2005) Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* 8, 1611–1618
- Rendell, L. *et al.* (2010) Why copy others? insights from the social learning strategies tournament. *Science* 328, 208–213
- Resnick *et al.* (2006) The value of reputation on eBay: a controlled experiment. *Exp. Econ.* 9, 79–101
- Livingston, J.A. (2005) How valuable is a good reputation? A sample selection model of internet auctions. *Rev. Econ. Statist.* 87, 453–465
- Stone, B. and Richtel, M. (2007) The hand that controls the sock puppet could get slapped. *N. Y. Times*, 16 July
- Tadelis, S. (2002) The market for reputations as an incentive mechanism. *J. Political Econom.* 110, 854–882
- Gambetta, D. (2009) *Codes of the Underworld: How Criminals Communicate*, Princeton University Press 336
- Pistone, D.J. and with Woodley, R. (1989) *Donnie Brasco.*, Signet
- Janlert, L.-E. (2006) Being and appearing: human interfaces in the digital age. In *Emerging Trends and Challenges in Information Technology Management* (Khosrow-Pour, M., ed.), pp. 232–234, IGI Publishing
- Esser, F. *et al.* (2000) Spin doctoring in British and German election campaigns –how the press is being confronted with a new quality of political PR. *Eur. J. Commun.* 15, 209–239
- Milinski, M. and Wedekind, C. (1998) Working memory constrains human cooperation in the Prisoner's Dilemma. *Proc. Natl. Acad. Sci. U. S. A.* 95, 13755–13758
- Stevens, J.R. and Hauser, M.D. (2004) Why be nice? Psychological constraints on the evolution of cooperation. *Trends Cogn. Sci.* 8, 60–65
- Oda, R. *et al.* (2009) Altruists are trusted based on non-verbal cues. *Biol. Lett.* 5, 752–754
- Stirrat, M. and Perrett, D.I. (2010) Valid facial cues to cooperation and trust: male facial width and trustworthiness. *Psychol. Sci.* 21, 349–354
- Trivers, R. (2000) The elements of a scientific theory of self-deception. *Ann. N. Y. Acad. Sci.* 907, 114–131
- Hoorens, V. (1993) Self-enhancement and superiority biases in social comparison. *Eur. Rev. Social Psychol.* 4, 113–139
- Valdesolo and DeSteno, D. (2007) Moral hypocrisy: social groups and the flexibility of virtue. *Psychol. Sci.* 18, 689–690
- Sommerfeld, R.D. *et al.* (2008) Multiple gossip statements and their effect on reputation and trustworthiness. *Proc. Biol. Sci.* 275, 2529–2536
- Heintz, C. (2006) Web search engines and distributed assessment systems. *Pragn. Cogn.* 14, 387–409
- Gürer, O. *et al.* (2006) The competitive advantage of sanctioning institutions. *Science* 312, 108–111
- Kreps, D.M. (1990) Corporate culture and economic theory. In *Perspectives on Positive Political Economy* (Alt, J.E. and Shepsle, K.A., eds), Cambridge University Press
- Suzuki, S. and Akiyama, E. (2005) Reputation and the evolution of cooperation in sizable groups. *Proc. Biol. Sci.* 272, 1373–1377

- 51 Chater, N. *et al.* (2008) A new consequence of Simpson's paradox: stable cooperation in one-shot prisoner's dilemma from populations of individualistic learners. *J. Ex Psychol. Gen.* 137, 403–421
- 52 Hagen, E.H. and Hammerstein (2006) Game theory and human evolution: A critique of some recent interpretations of experimental games. *Theoret. Pop. Biol.* 69, 339–348
- 53 Bateson, M. *et al.* (2006) Cues of being watched enhance cooperation in a real-world setting. *Biol. Lett.* 2, 412–414
- 54 Mifune, N. *et al.* (2010) Altruism toward in-group members as a reputation mechanism. *Evol. Hum. Behav.* 31, 109–117
- 55 Frank, R.H. (2006) Cooperation through moral commitment. in *Empathy and Fairness (Novartis Symposium 278)*, Foundation, N. ed.), Wiley
- 56 Zajonc, R.B. (1965) Social facilitation. *Science* 149, 269–274
- 57 Tennie, C. *et al.* (2009) The meat-scrap hypothesis: small quantities of meat may promote cooperative hunting in wild chimpanzees. *Behav. Ecol. Sociobiol.* 63, 421–431
- 58 Tomasello, M. (2009) *Why We Cooperate*, MIT Press, (Cambridge, MA)
- 59 Trivers, R.L. (1971) The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57
- 60 Nowak, M.A. and Sigmund, K. (2005) Evolution of indirect reciprocity. *Nature* 437, 1291–1298
- 61 Clutton-Brock, T. (2009) Cooperation between non-kin in animal societies. *Nature* 462, 51–57
- 62 Yamamoto, S. and Tanaka, M. (2009) Do chimpanzees spontaneously take turns in a reciprocal cooperation task? *J. Com. Psychol.* 123, 242–249
- 63 Warneken, F. *et al.* (2007) Spontaneous altruism by chimpanzees and young children. *PLoS Biol.* 5, e184
- 64 Melis, A. *et al.* (2008) Do chimpanzees reciprocate received favours? *Anim. Behav.* 76, 951–962
- 65 Grutter, A.S. and Bshary, R. (2003) Cleaner wrasse prefer client mucus: support for partner control mechanisms in cleaning interactions. *Proc. Biol. Sci.* 270 (Suppl 2), S242–244
- 66 Bshary, R. (2002) Biting cleaner fish use altruism to deceive image-scoring client reef fish. *Proc. Biol. Sci.* 269, 2087–2093
- 67 Bshary, R. and Grutter, A.S. (2006) Image scoring and cooperation in a cleaner fish mutualism. *Nature* 441, 975–978
- 68 Potts, G.W. (1973) Ethology of labroides-dimidiatus on aldabra. *Anim. Behav.* 21, 250–291
- 69 Russell, Y.I. *et al.* (2008) Image scoring in great apes. *Behav. Processes* 78, 108–111
- 70 Subiaul, F. *et al.* (2008) Do chimpanzees learn reputation by observation? Evidence from direct and indirect experience with generous and selfish strangers. *Anim. Cogn.* 11, 611–623
- 71 Silk, J.B. (2007) Social components of fitness in primate groups. *Science* 317, 1347–1351
- 72 Brosnan, S. *et al.* (2009) Chimpanzees do not develop contingent reciprocity in an experimental task. *Anim. Cogn.* 12, 587–597
- 73 Frith, C.D. and Frith, U. (2008) The self and its reputation in autism. *Neuron* 57, 331–332
- 74 Evans, C.S. and Marler (1991) On the use of video images as social-stimuli in birds - audience effects on alarm calling. *Anim. Behav.* 41, 17–26
- 75 Ziege, M. *et al.* (2009) Audience effects in the Atlantic molly -prudent male mate choice in response to perceived sperm competition risk? *Front. Zool.* E6, 17
- 76 Jones, S.S. *et al.* (1991) An audience effect on smile production in 10-month-old infants. *Psychol. Sci.* 2, 45–49