



REVIEW

No good deed goes unpunished: the social costs of prosocial behaviour

Nichola J. Raihani^{1*}  and Eleanor A. Power² 

¹Department of Experimental Psychology, University College London, 26 Bedford Way, London WC1H 0AP, UK and

²Department of Methodology, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK

*Corresponding author. E-mail: nicholaraihani@gmail.com

Abstract

Performing costly helpful behaviours can allow individuals to improve their reputation. Those who gain a good reputation are often preferred as interaction partners and are consequently better able to access support through cooperative relationships with others. However, investing in prosocial displays can sometimes yield social costs: excessively generous individuals risk losing their good reputation, and even being vilified, ostracised or antisocially punished. As a consequence, people frequently try to downplay their prosocial actions or hide them from others. In this review, we explore when and why investments in prosocial behaviour are likely to yield social costs. We propose two key features of interactions that make it more likely that generous individuals will incur social costs when: (a) observers infer that helpful behaviour is motivated by strategic or selfish motives; and (b) observers infer that helpful behaviour is detrimental to them. We describe how the cognition required to consider ulterior motives emerges over development and how these tendencies vary across cultures – and discuss how the potential for helpful actions to result in social costs might place boundaries on prosocial behaviour as well as limiting the contexts in which it might occur. We end by outlining the key avenues and priorities for future research.

Keywords: Prosocial behaviour; reputation; modesty; cooperation; partner choice; morality

Social media Summary: Good deeds can result in social costs, particularly when observers suspect calculated or malign motives.

Introduction

Investing in costly cooperative behaviour can yield benefits

Humans are outliers in the extent and frequency of our cooperation, particularly when this occurs with non-relatives and in settings where the potential for direct reciprocal benefits is low (Raihani & Bshary, 2015; Rand & Nowak, 2013; Melis & Semmann, 2010; Bshary & Raihani, 2017; Henrich, 2017; Henrich et al., 2005; Boyd, 2017). Helping behaviour may be under positive selection because these costly investments allow an actor to signal something about themselves to observers (Barker et al., 2019). For instance, helpful acts can allow an individual to reveal their type (Gintis et al., 2001) or their intentions (McNamara & Barta, 2020; Roberts, 2020; Singh & Hoffman, 2021). Receivers can benefit from attending to such signals as this allows them to identify and preferentially assort with cooperative interaction partners (Barclay, 2016; Baumard et al., 2013; McNamara et al., 2008). Theoretical models have confirmed the evolutionary logic of reputation-based cooperation, via either indirect reciprocity (Kandori, 1992; Ohtsuki & Iwasa, 2006) or reputation-based partner choice (Fu et al., 2008; Li, 2014; Roberts, 1998, 2015). Experimental evidence from laboratory studies

(including from non-humans; Bshary & Grutter, 2006) shows that the possibility of attracting partners is likely to be an important force for explaining the evolution and persistence of cooperation among non-relatives (Barclay & Barker, 2020; Barclay & Raihani, 2016; Raihani & Barclay, 2016; Sylwester & Roberts, 2010, 2013; Barclay & Willer, 2007; Tognetti et al., 2012; Van Vugt & Hardy, 2009; Fehrler & Przepiorka, 2016). Data from real-world settings (Hardy & Norgaard, 2016; Milinski et al., 2002a; Przepiorka et al., 2017; Raihani & Smith, 2015; Resnick et al., 2006; Yoeli et al., 2013) from many societies around the world (Bliege Bird et al., 2012; Bliege Bird & Power, 2015; Gurven et al., 2000; Lyle & Smith, 2014; Macfarlan et al., 2012, 2013; Macfarlan & Lyle, 2015; Power & Ready, 2018; Smith et al., 2003; Smith & Bird, 2000; von Rueden et al., 2019) also highlight the benefits that stem from being generous (or being seen to be generous): generous individuals are more likely to be helped themselves (Bliege Bird & Power, 2015; Gurven et al., 2000; Lyle & Smith, 2014; Macfarlan et al., 2012, 2013; Power & Ready, 2018), are healthier (Lyle & Smith, 2014) and have more and healthier offspring (von Rueden & Jaeggi, 2016).

But benefits do not always follow from prosocial actions

It is perhaps less well appreciated that cooperative actions can result in social costs rather than benefits. By revealing their cooperative intentions, individuals might leave themselves vulnerable to exploitation by others (e.g. Berezkei et al., 2015), particularly when players only expect to interact once (Roberts, 2020). Yet signalling prosociality can also invite hostility and competition from others. For instance, individuals often refuse offers of help (Ackerman & Kenrick, 2008), dislike those who offer help (Fisher et al., 1982) and may find overly generous individuals unattractive (Bhagal et al., 2020). Those who perform morally laudable behaviours in public also risk being accused of ‘virtue signalling’ (Levy, 2020). In laboratory social dilemma games, people’s antipathy towards prosocial individuals often manifests as antisocial punishment (Herrmann et al., 2008; Sylwester et al., 2013). Although people generally tend to direct punishment at the free-riders in the group (Falk et al., 2005; Fehr & Gächter, 2000, 2002; Johnson et al., 2009; Nikiforakis & Engelmann, 2011), a persistent fraction of players in social dilemma games sacrifice some of their personal endowment to fine cooperative co-players (Herrmann et al., 2008; Raihani & Bshary, 2019; Sylwester et al., 2013). Variants of such games that allow players to vote to evict members of the group rather than to financially punish them find similar results: the most cooperative group members are evicted as often as the worst free-riders (Parks & Stone, 2010). The tendency to dislike or to disparage prosocial or morally laudable others has also been studied under the banner of do-gooder derogation (Bai et al., 2019; Bolderdijk et al., 2018; Minson & Monin, 2012; Monin, 2007; Sparkman & Attari, 2020a; Zane et al., 2016) – a phenomenon whereby individuals who perform morally laudable actions (e.g. refraining from eating animal products or defending minority groups) are derogated by peers. This tendency to dislike generous or moral others has also been found in children as young as 8 years old (Tasimi et al., 2015). Although it is clear that good deeds can result in social costs, less work has systematically explored when and why these costs arise. We attempt to do so in this paper.

When and why might helpful signals yield social costs?

We propose that generosity is least likely to yield benefits and most likely to result in social costs in two, non-mutually exclusive scenarios:

- (1) observers infer that helpful behaviour is motivated by strategic or selfish motives;
- (2) observers infer that helpful behaviour is detrimental to them.

In what follows, we use the word observer to mean ‘someone who learns about the prosocial action of another individual’. In principle, the observer could learn because they are the recipient of the prosocial act or because they are an uninvolved bystander. We use the word observer to denote both possibilities. We expand on these predictions below.

Observers infer strategic or selfish motives

Praiseworthiness depends on attributing intentions to actors (Anderson et al., 2020; Bénabou & Tirole, 2006; Pizarro et al., 2003; Siegel et al., 2017), but the motives underpinning helpful actions are not visible to an observer. Instead, motives are typically inferred based on the helper's behaviour (e.g. the type and extent of help offered), any emotions revealed during the behaviour and the context (including the sociocultural setting, see Barrett et al., 2016; Henrich, 2020) in which the behaviour occurs.

Motives and outcomes do not always cohere: bad outcomes can stem from benign intentions, and good outcomes can stem from malign or selfish intentions. In principle, helpful actions can stem from proximately 'altruistic' motivations – that is, a genuine desire to help others (e.g. Andreoni, 1990; Batson et al., 1981; Bénabou & Tirole, 2006; Harbaugh et al., 2007). Nevertheless, at least part of the motive underpinning helpful actions may also be strategic or self-interested. For example, in experimental Dictator and Ultimatum games, a standard result is that people typically offer more money to a partner in the Ultimatum than in the Dictator game (Camerer, 2011). This result can be understood in light of the strategic nature of the Ultimatum game (where an offer that is perceived to be too low can be rejected by the responder) compared with the non-strategic nature of the Dictator game (where receivers must accept whatever share the dictator sends). Similarly, people often behave more generously in situations where their behaviour can be observed (e.g. Andreoni & Petrie, 2004; Bereczkei et al., 2007; Harbaugh, 1998; Hoffman et al., 1994; Yoeli et al., 2013) and can affect how others may treat them in future (Barclay & Willer, 2007; Milinski et al., 2002b; Semmann et al., 2004; Sylwester & Roberts, 2010, 2013), which also indicates that prosocial actions are influenced by strategic concerns.

These patterns hint at the complexity of decoding the motives underpinning generosity: some helpful behaviour may well be motivated by pure altruism, but it is clear that humans are also sensitive to the possibility of personal gains and adjust their investments in helpful behaviour accordingly (Semmann et al., 2004), even from a young age (Warneken et al., 2019). Observers (and recipients) of prosocial acts therefore face the difficult task of reconciling the actions they observe with a range of possible underlying motives. Some of these motives are likely to be more indicative of the helper's underlying cooperative disposition than others.

We propose that observers will be more likely to infer that help stems from strategic rather than altruistic motives when helpful actions appear to be aggrandising, for example if they are performed in public or are actively broadcast to others by the helpful individual. Several studies have shown that individuals who advertise their good deeds (Berman et al., 2015; De Freitas et al., 2019; Siem & Stürmer, 2018) or who know that their good deeds will be visible to others (Gambetta & Przepiorka, 2014) risk being perceived as less generous by observers (Srna et al., 2017). Similar results have been obtained in the context of corporate social responsibility (CSR). Specifically, consumers are more sceptical of CSR activities when they suspect that these deeds stem from profit-seeking rather than benevolent motives (Berge & Arshad, 2020; Chernev & Blair, 2015; Sen et al., 2016) and they downgrade their assessment of CSR initiatives when they learn of these from the company's advertising rather than from other sources (Yoon et al., 2006).

In general, positive acts are more likely to result in reputation benefits when observers can infer that the helper is motivated by empathic concern or that they are eager and willing to help. Indeed, the reputation consequences of actions may be affected more by the emotions attributed to the actor than to the outcome of the action itself (Yudkin et al., 2019). People rely more on emotions than on outcomes when judging helpful individuals because emotions provide a more reliable cue as to the person's underlying character or disposition (Barasch et al., 2014; Frank, 1991; Hoffman et al., 2015; Levine et al., 2018; Reed et al., 2012) or to the value they place on the relationship (Aktipis et al., 2018; Ames et al., 2004; Delton & Robertson, 2016; Frank, 1991; Hoffman et al., 2015) – and are therefore a more reliable guide to how the person might behave in future (see also Hirshleifer, 1987; Singh & Hoffman, 2021). For instance, people who perform prosocial actions but display

reluctance or negative emotions while doing so are unlikely to accrue reputation benefits as the negative emotional signal nullifies the positive act (Ames & Johar, 2009; Carlson & Zaki, 2018; Krull et al., 2008). Similarly, people who express mixed motives for helping or who help without expressing empathy are also less positively evaluated by others (Erlandsson et al., 2020). When evaluating the moral consequences of charitable donations, people rate those who donate time as being more praiseworthy than those who donate money (even though people also acknowledge that time donations are often less efficient; Johnson & Park, 2019). This evaluative premium on time arises in part because people perceive time donations as signalling a greater emotional investment in the cause and therefore a better indicator of the donor's underlying character.

The idea that observers try to discern a helper's underlying character or disposition when evaluating prosocial acts is also evident in the negative reaction to hypocrisy or inconsistent behaviour (Effron et al., 2018; Jordan et al., 2017). For example, in one vignette-based study, individuals who signalled moral virtue (by criticising others) but subsequently behaved immorally themselves were more negatively judged than individuals who admitted their moral failings (Jordan et al., 2017). Similar reactions are observed in the domain of CSR and can help to explain why these activities sometimes backfire: if companies with bad reputations engage in CSR, consumers are more likely to infer that these socially responsible actions stem from a desire to whitewash a bad reputation (Yoon et al., 2006). CSR activities that seem at odds with the broader mission of the company can also negatively impact brand evaluations (Silver et al., 2021; Torelli et al., 2012). Indeed, companies that emphasise the business case for CSR tend to be more positively evaluated than those who emphasise the moral case because emphasising the business case allows the company to avoid accusations of moral hypocrisy (Hafenbrädl & Waeger, 2019). A key point here seems to be that being upfront about the underlying reasons (or potential benefits) associated with performing prosocial deeds can quell the suspicion these deeds might otherwise arouse in some cases.

Relatedly, observers might also suspect that helping stems from strategic rather than altruistic motives if the helper deliberates or assesses the costs before deciding whether to help (Hoffman et al., 2015). Participants who cooperate without knowing the cost of doing so (Jordan et al., 2016) – or who make their decision quickly (Critcher et al., 2013; Evans & van de Calseyde, 2017) – are perceived as more prosocial or trustworthy by others – as are those who make deontological, rather than utilitarian, decisions when faced with moral dilemmas (Everett et al., 2016). Similarly, individuals who donate based on the cost-effectiveness of the charity, rather than based on the degree of empathy they feel for the beneficiaries, are evaluated less positively by others (Monteleagre et al., 2020). Subjects who cooperate in seemingly one-shot encounters are also perceived as more trustworthy than those who cooperate in encounters that they know will be repeated (Johnsen & Kvaløy, 2016), perhaps because cooperating in interactions that are known to be repeated is more indicative of strategic motives.

The appearance of strategic, calculating decision-making is particularly costly in long-term reciprocal exchanges ('friendships'), when an effort to quickly reciprocate and erase a debt may reduce what could be seen as an open-ended cooperative partnership to a disinterested series of transactions (Bourdieu, 1990; Silk, 2003). The logic of tit-for-tat in a long-term relationship can be damaging because it implies low interdependence (Aktipis et al., 2018; Kim et al., 2019; Roberts, 2005). This can explain why subjects tend to repay favours from strangers more quickly than those from friends (Boster et al., 1995) and can take offence – and judge friendships as less close – when friends immediately reciprocate kind acts (Shackelford & Buss, 1996; Silk, 2003) or give gifts with the expectation of adulation or return favours (Lee, 1969).

Motives might also be inferred based on whether the helper materially benefits from their actions. For example, donors to charity are perceived as less charitable if they donate to a cause to which they have a personal connection (Lin-Healy & Small, 2012), and advocates for a charitable cause are perceived as less sincere and garner fewer donations for the cause if the advocates reap financial benefits from this association (Barasch et al., 2016). In corporate settings, organisations that perform good deeds are given less credit if doing good helps their bottom line (Makov & Newman, 2016;

Newman & Cain, 2014), an effect known as ‘tainted altruism’ (Newman & Cain, 2014). In keeping with this view, people more positively judge helpful actions when these involve a degree of personal sacrifice (Bigman & Tamir, 2016; Carlson & Zaki, 2018; Johnson, 2018; Kawamura et al., 2020; Nelissen, 2008; Visserman et al., 2018). Indeed, in the context of charitable donations, people are more sensitive to personal sacrifice than social benefit when evaluating prosocial behaviour (Johnson, 2018). This increased perception of costs, and so an assumed lesser material benefit, can explain why lower status helpers are sometimes perceived in a more positive light. A vignette study found that lower class donors (measured in terms of occupation and salary) were more likely to be seen as prosocial, compared with higher class donors, when giving the same (either relative or absolute) donation to a charity, as the lower-status individuals were seen as more authentically motivated (Yuan et al., 2018).

Observers infer that helping behaviour is detrimental to them

Although prosocial acts are typically interpreted as yielding benefits to others, this may not always be the case. Instead, observers might sometimes perceive (accurately or not) that a prosocial act is likely to harm, rather than to benefit, them. For this reason, the reputation consequences of helping are also likely to depend on who is watching: helpful acts might improve your reputation in the eyes of some individuals but not others. Broadly, it seems plausible that observers who incur costs or infer that they are disadvantaged as a consequence of the helpful actions of another individual will be more likely to evaluate the helpful individual negatively. We outline some specific cases below.

One of the most obvious ways that observers might incur costs from the actions of helpful individuals is due to social comparison. A good reputation is, by definition, a positional good – a person’s reputation is ‘good’ in relation to the reputations of other individuals to whom that individual is compared (Barclay, 2011, 2013, 2016; Samu et al., 2020). Prosocial actions that improve one person’s reputation (or can be construed as potentially doing so) can therefore provoke competitive responses from those whose reputation may suffer by comparison (e.g. Herrmann et al., 2019; Macfarlan et al., 2012; McAndrew & Perilloux, 2012; Pleasant & Barclay, 2018; Raihani & Smith, 2015; Sylwester & Roberts, 2013). Similarly, if status hierarchies are formed in part on the basis of patronage and largesse, then one person’s generous acts may have the effect of lowering others’ relative positions within that hierarchy. Gift-giving or hosting large feasts, for example, may be arenas for status competition and the demonstration of power and authority (Boone, 1998); such acts may therefore invite negative reactions from those displeased with how they will look in comparison. Studies exploring intergroup relations have shown that help directed from higher to lower status groups can be resisted as it can be interpreted as an attempt to assert or maintain dominance. For example, when asked to imagine receiving help from an out-group, Israeli Arabs were more likely to judge help from Israeli Jews as being directed at reinforcing existing power dynamics (Halabi et al., 2016; Nadler & Halabi, 2006). Cross-cultural work using economic games also suggests that people resist generous overtures that can be interpreted as attempts to establish dominance (Henrich et al., 2005).

Antisocial punishment and other hostile reactions to cooperators observed in social dilemmas are thought to be prompted at least in part by social comparison. For example, in a public goods game, where players could vote to evict one player from the group after each round, players often voted to evict the most generous member from the group (Parks & Stone, 2010). When asked to justify these decisions, participants frequently referenced their own reputation, for example ‘No one else is doing what he does. He makes us all look bad’ and ‘Next to Person Blue I look like a bad guy, but I don’t think I am’. Other studies have also shown that excessive generosity is judged unfavourably and may even be punished (Bahry & Wilson, 2006; Irwin & Horne, 2013; Kawamura & Kusumi, 2020). In contrast, antisocial punishment of overly cooperative individuals is rare when the punisher is a third party, rather than a player who was involved in the public goods game (Bone et al., 2014). Since the third party cannot contribute to the public good, their reputation as a generous or cooperative individual is not affected by the investments of the other players. The fact that uninvolved parties do not tend

to antisocially punish therefore supports the idea that hostile reactions to helpful individuals are mediated by social comparison. In a more recent study, Pleasant and Barclay (2018) showed that antisocial punishment in a public goods game was most prevalent when individuals were being evaluated as a partner for a subsequent Trust Game, which further supports the idea that antisocial punishment operates via concern for reputation and relative standing. As with antisocial punishment, experimental studies have shown that do-gooder derogation frequently stems from competitive motives – because people who appear to be morally superior can prompt less moral individuals to evaluate themselves more negatively (Minson & Monin, 2012; Monin, 2007; Monin et al., 2008).

Observers might also negatively evaluate helpful individuals when the helper's and the observer's goals are misaligned (Melnikoff & Bailey, 2018). For example, people prefer moral traits in others (such as mercifulness, loyalty and trustworthiness) only when these traits are conducive to their current goal. In scenarios where it would benefit these individuals to interact with merciless, fickle or dishonest others, then these traits are more positively evaluated (Melnikoff & Bailey, 2018). A similar discrepancy has been observed in moral dilemmas, such as the trolley problem, where people prefer deontological agents when choosing social or romantic partners but prefer consequentialist (utilitarian) individuals when electing political leaders (Everett et al., 2018).

Another way in which a helper's goals might be misaligned with an observer's is when the help offered to one person is traded off against the help that can be offered to another (Law et al., 2019). For example, in Western cultures, helping strangers is typically perceived as being a morally positive thing to do (Henrich, 2020; McManus et al., 2020). However, even in a sample of US residents, people who help strangers instead of helping family were generally judged more negatively (McManus et al., 2020), indicating that helpers' reputations are formed on the basis of who is helped and the trade-offs involved. Related to the above, helpful acts can sometimes threaten the observer's existing relationship with the helper – for example by calling into question the closeness of the relationship (Law et al., 2019) and thereby potentially sparking jealousy (Krems et al., 2021).

Helpers might also be judged more negatively when the help disadvantages the observer's in-group or is directed towards the observer's out-group (cf. Fessler et al., 2015). For example, one study measured participants' reactions to another individual who confronted discrimination, involving both sexism and racism as a function of whether the confronter was part of the disadvantaged group (e.g. the confronter is also a woman in the sexism example) or part of the advantaged group (e.g. the confronter is a man in the sexism example; Kutlaca et al., 2020). In this study, participants were more likely to judge that the confronter had over-reacted if both the participant and the confronter were members of the same advantaged in-group. Similarly, whistleblowers within organisations are seldom rewarded for their actions and are often more likely to be bullied than those who keep quiet (Bjørkelo et al., 2011). The evaluation of any particular 'helpful act' must, then, lie in the eye of the beholder.

Reputation management strategies in humans

Given the potential social costs associated with performing overtly generous or moral actions, it is not surprising that people manage their reputations in quite sophisticated ways, for example by refraining from doing good deeds in public or by actively hiding their good deeds from others. While many studies have shown that people are more cooperative when their behaviour is public (e.g. Andreoni & Petrie, 2004; Bekkers & Wiepking, 2011; Bereczkei et al., 2007; Bohnet & Frey, 1999; Piazza & Bering, 2008; Yoeli et al., 2013), other studies report mixed or contrasting results. Notably, a recent meta-analysis found that although there is a small and statistically significant positive effect of observability on prosocial behaviour, the effect is highly heterogeneous and context-dependent (Bradley et al., 2018). For example, public recognition can decrease people's willingness to donate blood (Shi, 2011) or to donate to charity (Denis et al., 2020; Savary & Goldsmith, 2020; Simpson et al., 2018). Experimental studies have shown that individuals feel uncomfortable when they learn that they are morally superior to their peers (Cowgill, 2017) or when their moral actions are highlighted to others (Wang & Tong, 2015) and these negative feelings might prompt individuals to hide their

good deeds from others. For instance, when making charitable donations online, people often hide their donations from others by making them anonymously (*Big Charitable Gifts*, 2020; Imada, 2020) and the tendency to donate anonymously (either by hiding their name or the amount donated) is higher when the donation is especially generous (Mokos & Scheuring, 2019; Peacey & Sanders, 2014; Raihani, 2014). Similar results have been obtained in experimental settings (Jones & Linardi, 2014), although we note that most of these studies involve WEIRD (Henrich et al., 2010) populations and it remains an open question as to whether these patterns would generalise to different cultures.

People also seem to be sensitive to the possibility that their good deeds will reveal self-interested intentions, and often work to prevent such inferences from being made. For example, when donating to charity, people are more likely to perform charitable actions that are perceived as being painful or difficult (Olivola & Shafir, 2013), perhaps to lessen any suspicions of self-interest in ostensibly altruistic behaviour. Similarly, time donations are viewed as more morally praiseworthy than money donations and – as we might expect – people often prefer to donate time than money, even when the latter could have a larger impact (Johnson & Park, 2019). In studies where people could cooperate in a calculated or uncalculated manner (by revealing the cost associated with cooperating), people are more likely to opt for uncalculated cooperation when they know their decision is observed by others (Jordan et al., 2016). Similarly, hunters, especially of large game, may practice ‘pecuniary distancing’ by actively avoiding having control over the divvying up and sharing of their catch, so that others cannot infer selfish intent (Bliege Bird & Power, 2015).

Individuals also manage their prosocial reputation in other ways, for example by selectively investing in subtle, rather than ostentatious signals (Bliege Bird et al., 2018). The very subtlety and discretion of acts may help ensure that they are seen as generous and well intentioned: when no audience is there to provide a reputational boon, it should be much harder to infer that the intention was one of self-aggrandisement. Helpful acts which are experienced and observed only (or primarily) by the intended recipient may then do much to assure that person of the giver’s commitment to their continued relationship.

Recent theoretical work has also shown how obscuring or ‘burying’ costly signals can also be under positive selection (Hoffman et al., 2018). This occurs because buried signals come with a tag that identifies them as having been buried (e.g. we might know whether our wealthy friend chose to donate anonymously on a fundraising platform). This allows the fact that the signal has been buried to act as a signal in itself. Although initial work found that only the ‘best’ signallers would evolve to bury their signals, a subsequent analysis found that a more universal modesty norm can also spread through populations (Quillien, 2019). Empirical studies support these models, by showing that modesty begets modesty and that modest actions are likely to be socially learned. For example, in one study of online charitable giving, the presence of anonymous donations encouraged subsequent donors to anonymise their own donations (Burtch et al., 2016).

In general, we suggest that concern for reputation might often manifest as behaviour aimed at preventing reputational damage, rather than behaviour aimed at actively improving reputation. This is partly because, as the evidence above indicates, prosocial acts do not always result in social benefits. Not only that, but there is a further asymmetry in the social consequences of blameworthy and praiseworthy acts: there is often more to be lost from being seen as blameworthy than to be gained from being seen as praiseworthy (reviewed in Anderson et al., 2020). For example, breaking a promise is viewed as morally contemptible, but over-delivering on a promise carries no reputation benefits relative to simply keeping it (Gneezy & Epley, 2014). Similarly, moral judgements of selfish actions are moderated by the magnitude of the selfishness (more selfish actions are judged morally worse), but the converse is not true for selfless actions: the benefits of behaving prosocially do not scale linearly with the magnitude of the prosocial deeds (Klein et al., 2015; Klein & Epley, 2014). Other work has shown that blameworthiness judgements are often more extreme than praise judgements (Guglielmo & Malle, 2019) and more sensitive to outcomes. Given all this, the risk of a generous act resulting in social costs may weigh more heavily in people’s minds than the potential benefits that could stem from it.

The cognitive underpinnings of managing and evaluating reputations

There is clearly more to reputation management – and impression formation – than the simple performance of prosocial actions to reap social benefits. Instead, when forming an impression of others, we have to be able to infer the motives and intentions that accompany positive actions. When managing our own reputation, we also have to be able to infer how our own actions – and potentially divergent perceptions of our intentions – will affect how others form and update their assessments of us. The ability to manage one's own reputation and to evaluate the reputations of others relies on 'epistemic vigilance' (Sperber et al., 2010) – a suite of cognitive abilities which probably evolved to help us discern reliable from unreliable communication. These abilities emerge at different stages of development (Banerjee, 2000; Banerjee et al., 2020; Engelmann & Rapp, 2018; Silver & Shaw, 2018; Sperber et al., 2010) and show differential expression in reputation-based cooperation across societies (Amir & McAuliffe, 2020; Cowell et al., 2017).

A key feature of epistemic vigilance is the ability to infer intentions that appear to be at odds with observed actions (e.g. *The actions suggest X but the motive is Y*). This form of intention attribution is distinct from and computationally more demanding than the simpler version of intention attribution (likely to be present in non-human apes and very young children; Tomasello et al., 2005), where observers impute intentions that cohere with a target's observed behaviour. The development of scepticism – where we suspect that an agent is trying to present themselves in a favourable light – involves doing something even more sophisticated, however: we may intuit not only that A's intentions are at odds with their actions, but also that A would like us (or other observers) to form a false belief about their intentions (e.g. *A wants me to believe their motive is X when I suspect in fact it is Y*). This latter awareness involves reasoning at a higher meta-representational level (Sperber et al., 2010), which may account for its relatively late emergence during childhood (Heyman et al., 2007; Heyman et al., 2014).

Children start to manage their own reputation – and to evaluate the reputations of others – in early childhood (Banerjee et al., 2020; Engelmann & Rapp, 2018; Silver & Shaw, 2018). The ability to form impressions of others seems to emerge earlier in development than the ability to manage one's own reputation: indeed, some studies have even found that pre-verbal infants evaluate others on the basis of helpfulness (Hamlin, 2013; Hamlin et al., 2011; but see Salvadori et al., 2015; Schlingloff et al., 2020 for failed replications). Whether or not the ability to evaluate others emerges in preverbal infants, this ability does seem to be reliably evident by early childhood. By the time they are around two and a half years old, children can discriminate a nice partner from a mean one, only on the basis of observation, and prefer to interact with the nice partner (Herrmann et al., 2013). Children as young as three years old expect that individuals will share information that presents them in a good light, for example, by revealing their positive performance on a puzzle but withholding their failures (Hicks et al., 2015); and by the age of five, children begin to gossip, offering their opinion on the character of others to help a third party to decide who to interact with (Engelmann et al., 2016). Four-year-olds appreciate that others will lie to benefit themselves and, by the time they are six, children distinguish between self-interested lies and lies intended to benefit others (Banerjee et al., 2020; Mills & Keil, 2005). In one study, children were asked to evaluate a protagonist who either plagiarised someone else's story (and took the credit for themselves) or who lied to give the credit for their own story to another child (Shaw & Olson, 2015); children were more likely to evaluate the person who stole the credit negatively – and to evaluate the person who gave credit away positively.

In contrast to the early emergence of impression formation, children do not begin managing their own reputation until they are around five years old. By this age, children (but not chimpanzees) are sensitive to the presence of an audience when making prosocial decisions (Engelmann et al., 2012) – and give more when their decisions will be revealed to others (Leimgruber et al., 2012). Nevertheless, before middle childhood (around the age of 8) children tend not to reason explicitly about how their actions will affect other individuals' impressions; rather, reputation management in young children has been referred to as 'implicit' and may overlap with what has been called 'audience effects' in non-human species. By middle childhood, children start to explicitly reason about how their

actions will affect other people's impressions of them (Banerjee et al., 2020; Engelmann & Rapp, 2018) – and it is at this point that children also start to understand that other individuals may do the same (Heyman et al., 2007, 2016; Heyman et al., 2014). For example, although young children do not differentiate between acts of generosity that are performed in public vs. those that are done in private, by the time they are eight years old, children reliably rate private acts of beneficence as more morally praiseworthy (Heyman et al., 2016; Heyman et al., 2014). In middle childhood, children also begin to exhibit scepticism about self-serving claims of prosociality (Amemiya et al., 2020; Heyman et al., 2007; Heyman et al., 2014), although this ability continues to develop over this period. For example, older children (aged nine to 10) appreciate that self-promotional statements are more morally permissible when they can also be construed as implicit offers of help (Amemiya et al., 2020).

At the same time as children start to appreciate that others strategically manage their reputation – and develop a sense of scepticism about people who use self-promotional strategies – they also begin to understand the benefits of modest actions and to evaluate modest others as more morally praiseworthy (Banerjee, 2000). The importance of modesty – and the extent to which children both strive to appear modest and approve of modesty in others – varies cross-culturally, being more strongly emphasised in East-Asian than in Western societies. For example, in a study involving Japanese and American schoolchildren (aged seven to 11 years), the tendency to approve of modest individuals (who lied about their prosocial actions) increased with age in both groups (Heyman et al., 2011). Nevertheless, the Japanese children were also more likely to disapprove of those who truthfully accepted the praise for their good deed, a pattern that was not observed among the American children. A more recent study comparing self-effacement in a sample of Chinese and Canadian children of the same age found that the tendency to eschew moral praise increases with age – but that this effect was more pronounced among the Chinese children (Fu et al., 2016). These and related studies strongly suggest a role for social learning in the development of modesty traits (see also Banerjee, 2000) and emphasise the fact that modesty norms (and associated reputation management strategies) are likely to vary cross-culturally (Genyue et al., 2011; Kanagawa et al., 2001; Yamagishi et al., 2012). Cross-cultural evidence generally suggests that responsiveness to local norms solidifies in middle childhood, and that this can precipitate important variation in prosocial behaviour (Amir & McAuliffe, 2020; Blake et al., 2015; Cowell et al., 2017; House et al., 2013, 2020a, b).

In sum, several cognitive abilities are implicated in even basic reputational models of cooperation – including the ability to mind-read and attribute intentions and motives to others, to be sensitive to the local norms in order to judge whether behaviours are praiseworthy or contemptible, and to infer how others might perceive us based on our actions and their own perceptions of our hidden intentions. These skills do not all emerge at a single set point in development, but are instead honed over childhood and into adolescence, are responsive to local cultural norms, and are generally highly contingent and situational.

Larger-scale consequences of perverse reputational incentives

Considering the social costs incurred by prosocial individuals and organisations has broader implications for the emergence and persistence of cooperation within groups, as well as for the effectiveness of these altruistic efforts. Many of the social interactions we face can be broadly described as social dilemmas: mutual cooperation yields the best collective payoffs, but individuals can derive a larger short-term benefit from acting selfishly. Finding ways to resolve such social dilemmas – and to encourage mutual cooperation – is both theoretically and practically important (e.g. Raihani & Aitken, 2011; Ostrom et al., 1999; Van Vugt, 2009; Penn, 2003). A key insight from experimental economics and social psychology is that many people are conditional cooperators (Fischbacher et al., 2001). Under this view, good deeds can spark further cooperation (Gino et al., 2009; Jordan et al., 2013; McAuliffe et al., 2015; Pfeiffer et al., 2005; Raihani & Hart, 2010) and uncooperative acts prompt reciprocal defection, potentially leading to the unravelling of cooperation (e.g. Raihani & Hart, 2010). Although we do not want to challenge the notion of conditional cooperation, we also want to

emphasise that cooperative individuals can also sometimes undermine the emergence of large-scale cooperation. For example, people often hold negative stereotypes of socially conscious individuals – and these perceptions can dissuade others from cooperating in these domains (Bashir et al., 2013; Fruhen & Flin, 2015). The investments of prosocial individuals can also deter individuals who might otherwise cooperate, either because they do not feel the same degree of commitment to the cause as these moral minorities (Kurz et al., 2020) or because they fear moral reproach if their own efforts fall short by comparison (Howe & Monin, 2017; Sparkman & Attari, 2020b; Stouten et al., 2013).

Moreover, negatively evaluating prosocial individuals and companies can act as a further obstacle to achieving large-scale cooperation because these social costs act as an additional disincentive to pay the costs associated with cooperative behaviour. For example, people and companies whose actions yield personal *and* social benefits are typically evaluated more negatively than counterparts whose actions yield no social benefits at all (Makov & Newman, 2016; Newman & Cain, 2014). By disparaging people who personally benefit from performing good deeds – or by enacting preferences against for-profit, for-good companies – people might disincentivise those who might otherwise invest in socially beneficial activities.

The special credit we afford to deeds seemingly motivated by ‘pure’ emotions compared with those that stem from ‘deliberative’ calculation can also have unwanted large-scale consequences. For instance, despite the large sums that are collectively donated to charity (around \$400 billion per year in the US), many people do not prioritise ‘effectiveness’ in the decision of which charity to support (Caviola et al., 2020). Instead, people support charities based on their subjective preferences (Berman et al., 2018), being more likely to donate to disaster-relief campaigns than to charities that aim to resolve recurring problems (Caviola et al., 2020) or donating to causes that are associated with single identifiable victims rather than those which will help larger numbers (Jenni & Loewenstein, 1997). Ineffective altruism can nevertheless make sense in light of its reputation consequences. For example, those who deliberate over which charity to support are evaluated less positively than those who use empathy as their moral compass (Montealegre et al., 2020) and, despite the fact that donating to charities in lower-income countries is more effective than donating to more local causes (Caviola et al., 2020), people often disapprove of those who forego opportunities to help at a local scale in favour of helping more distant others (Law et al., 2019; McManus et al., 2020). More generally, the social rewards that flow from prosocial acts tend to be sensitive to the costs and observability of the good deeds and typically do not scale with the efficacy of the helpful act, which can explain why people engage in ineffective helping (Borum et al., 2020).

The tendency to seek ulterior motives behind ostensibly prosocial actions can also have effects on self-assessments of our own helpful deeds, which subsequently affect the tendency that we will repeat them. Specifically, it appears to be easier to maintain a positive self-image and to derive a warm glow from giving when we can be assured that our own motives were not corrupted by extrinsic incentives (Gneezy et al., 2012; Savary et al., 2020; Savary & Goldsmith, 2020; Small & Cryder, 2016). It is now well established that external incentives for prosocial behaviour – such as thank you gifts or financial incentives, like tax breaks – frequently reduce people’s willingness to perform these prosocial acts (Chao, 2017; Gneezy & Rustichini, 2000; Mellström & Johannesson, 2008; Moussaoui et al., 2019; Newman & Jeremy Shen, 2012; Savary et al., 2020; Shi, 2011). These negative effects can be ameliorated by giving people the opportunity to forego the external incentive (Mellström & Johannesson, 2008) – and other work indicates that people often engage in such ‘motivational laundering’: foregoing an external incentive to remove ambiguity about the motives underlying their charitable actions (Kirgios et al., 2020).

Conclusions and future directions

In this final section, we outline some specific directions for future research, focusing on research gaps and open questions. One key hypothesis stemming from this review is that people might often seek to avoid social costs rather than to garner social benefits – particularly when it comes to reputation

management in daily life. Predictions stemming from this hypothesis cannot easily be tested with experimental paradigms that constrain the ways in which individuals are allowed to manage their reputation (for example, using experiments that only allow individuals to signal prosociality to attract interaction partners). Instead, we believe that a more naturalistic, observational approach as used in studies of animal behaviour would be helpful to address this hypothesis. The rich insights from ethnography and participant observation will surely have a role here. Alongside this, new technologies allow for experience sampling methods and have been used to great effect to address other aspects of social behaviour in daily life (e.g. Molho et al., 2020; Weiss et al., 2020). Vignette studies, wherein individuals are presented with a situation and give their reaction or outline potential courses of action may also provide a promising avenue (cf. Mathew & Boyd, 2014). Such methods could be used to measure the various ways that people manage their reputations and form and update impressions of others and to test the idea that avoiding social costs takes precedence over seeking reputation benefits.

Another key theme coming out of the existing literature on reputation-based cooperation is that both reputation management and impression formation rely on a suite of cognitive abilities, which seem to emerge during middle childhood. This is now a burgeoning area of research and great strides are being made in identifying the cognitive abilities that scaffold reputational strategies. However, it is still unclear which cognitive mechanisms are key to these abilities and how exactly these abilities are attained. Specifically, it is not clear whether the cognition supporting sophisticated forms of reputation-based cooperation emerges at relatively fixed points in development or whether the abilities have to be culturally learned and practised during childhood (Heyes, 2019; Luhrmann, 2011). We also know rather little about how individual variation in socio-cognitive abilities affects reputational strategies. For example, do individuals vary in their ability to both manage their own reputation and to form accurate impressions of others; and, if so, where does this variation come from? It may also be fruitful to explore metacognitive accuracy – specifically whether individuals are able to judge how signalling their prosocial deeds might be interpreted by others, as well as the factors that predict performance in this domain. We also note that previous studies have documented societal variation in the tendency to account for mental states when judging harmful outcomes (Barrett et al., 2016; Curtin et al., 2020): in countries that have stronger kin-based institutions, people are less likely to take an actor's mental state into consideration when forming moral judgements. As we have argued, intention attribution is also pivotal in determining how people respond to actions with beneficial outcomes but much of this evidence is derived from Western societies. A clear research priority is therefore to determine whether consideration of mental states when judging beneficial acts varies in a similar manner around the world.

It is also apparent that the content of the norms we use to guide our behaviour and form impressions of others varies hugely around the world. There is a clear need to understand how cognition and culture exert their independent and interacting effects on reputational strategies in humans; and this field would benefit from theoretically informed work addressing how and why such variation arises. For example, although it seems apparent from the studies above that modesty norms are more strongly endorsed in East Asian (compared with Western) societies, it is not especially clear why this is the case: the observation that modesty is more highly prized in some societies as compared with others falls short as an explanation as it simply restates the puzzle.

Crucial to any explanation of cross-cultural variation in the importance of reputation and social costs/benefits will be an understanding of the nature and structure of social relations: for example, the extent of people's interdependence (Aktipis et al., 2018; Roberts, 2005), their embeddedness in dense kin networks (Enke, 2019; Henrich, 2020; Schulz et al., 2019), their ability to extricate themselves from existing relationships (Jackson et al., 2012) and the importance of social relations to future generations (Bowles et al., 2010; Borgerhoff Mulder et al., 2009). Such variation in the form and salience of social relations, which can stem from differences in ecology and mode of production (Bowles et al., 2010; Lamba & Mace, 2011), as well as historical contingency and cultural norms, should help account for variations in social norms and institutions around the world. Comparative, cross-cultural

work that captures these aspects of social structure (whether through social network collection or field experiments that retain elements of people's social network position and connections) is needed to help us understand these points.

Finally, just as there is a need to explain variation across societies, so too is there a need to explain variation in the strategies of reputation management and impression formation used *within* societies. If the reaction to any prosocial act is contingent and context-specific, then we can expect that people's exposure to and concern for social costs will vary. An individual's social standing may have particular bearing on how prosocial acts are interpreted (and so, on the individual's propensity to carry them out). However, it is not immediately clear *how* this might feed in. For example, we may expect people of lower social standing to be viewed in a positive light because of the greater relative cost they incur for any helpful act (e.g. Yuan et al., 2018). Or, we could expect the opposite, with people of lower social standing being viewed with more scepticism: if they are more in need of the potential social benefits that stem from helpful acts, then they may more readily be assumed to have strategic aims, as compared with better off individuals who could be doing things not out of need, but out of 'the goodness of their heart'. Vignette studies could extend earlier work (Barrett et al., 2016) to determine whether and, if so, how judgments are dependent on the identity and status(es) of the actor involved. If there are indeed systematic biases in who faces derision for their prosocial acts and who garners praise, then it is also worth investigating the aggregate consequences of these inequalities. Formal and agent-based models may be particularly useful in studying the structural consequences of particular patterns of biases in interpretation.

To sum up, prosocial acts commonly yield reputation benefits, which can be advantageous in the context of partner choice. Nevertheless, there is also the potential for these investments to result in social costs – and this is most likely when observers can infer self-interested intentions or when observers believe that helpful actions have the potential to harm them. We have emphasised the contingent and context-specific nature of such evaluations. There is a clear need for more cross-cultural work in this field so that we address the extent to which some of the most frequently observed patterns are culturally specific or are more likely to be human universals.

Author contributions. Both authors contributed equally to this paper.

Financial support. NR is supported by a Royal Society University Research Fellowship and by the Leverhulme Trust.

Conflict of interests. NR is the author of the forthcoming trade book, *The Social Instinct: How Cooperation Shaped the World*.

Data availability. There are no data to report.

References

- Ackerman, J. M., & Kenrick, D. T. (2008). The costs of benefits: Help-refusals highlight key trade-offs of social life. *Personality and Social Psychology Review*, 12(2), 118–140. <https://doi.org/10.1177/1088868308315700>
- Aktipis, A., Cronk, L., Alcock, J., Ayers, J. D., Baci, C., Balliet, D., ... Winfrey, P. (2018). Understanding cooperation through fitness interdependence. *Nature Human Behaviour*, 2(7), 429–431. <https://doi.org/10.1038/s41562-018-0378-4>
- Amemiya, J., Liu, Z., Compton, B. J., & Heyman, G. D. (2020). Children's judgements of positive claims people make about themselves. *Infant and Child Development*, e2212. <https://doi.org/10.1002/icd.2212>
- Ames, D. R., Flynn, F. J., & Weber, E. U. (2004). It's the thought that counts: On perceiving how helpers decide to lend a hand. *Personality and Social Psychology Bulletin*, 30(4), 461–474. <https://doi.org/10.1177/0146167203261890>
- Ames, D. R., & Johar, G. V. (2009). I'll know what you're like when i see how you feel: How and when affective displays influence behavior-based impressions. *Psychological Science*, 20(5), 586–593. <https://doi.org/10.1111/j.1467-9280.2009.02330.x>
- Amir, D., & McAuliffe, K. (2020). Cross-cultural, developmental psychology: Integrating approaches and key insights. *Evolution and Human Behavior*, 41(5), 430–444. <https://doi.org/10.1016/j.evolhumbehav.2020.06.006>
- Anderson, R. A., Crockett, M. J., & Pizarro, D. A. (2020). A theory of moral praise. *Trends in Cognitive Sciences*, 24(9), 694–703. <https://doi.org/10.1016/j.tics.2020.06.008>
- Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, 100(401), 464. <https://doi.org/10.2307/2234133>

- Andreoni, J., & Petrie, R. (2004). Public goods experiments without confidentiality: A glimpse into fund-raising. *Journal of Public Economics*, 88(7), 1605–1623. [https://doi.org/10.1016/S0047-2727\(03\)00040-9](https://doi.org/10.1016/S0047-2727(03)00040-9)
- Bahry, D. L., & Wilson, R. K. (2006). Confusion or fairness in the field? Rejections in the ultimatum game under the strategy method. *Journal of Economic Behavior & Organization*, 60(1), 37–54. <https://doi.org/10.1016/j.jebo.2004.07.005>
- Bai, F., Wu, W., & Bao, S. (2019). Moral but dominant: When do-gooders get derogated. *Academy of Management Proceedings*, 2019(1), 14563. <https://doi.org/10.5465/AMBPP.2019.14563abstract>
- Banerjee, R. (2000). The development of an understanding of modesty. *British Journal of Developmental Psychology*, 18(4), 499–517. <https://doi.org/10.1348/026151000165823>
- Banerjee, R., Heyman, G. D., & Lee, K. (2020). The development of moral self-presentation. In L. A. Jensen (Ed.), *The Oxford handbook of moral development* (pp. 91–109). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190676049.013.6>
- Barasch, A., Berman, J. Z., & Small, D. A. (2016). When payment undermines the pitch: On the persuasiveness of pure motives in fund-raising. *Psychological Science*, 27(10), 1388–1397. <https://doi.org/10.1177/09567976166638841>
- Barasch, A., Levine, E. E., Berman, J. Z., & Small, D. A. (2014). Selfish or selfless? On the signal value of emotion in altruistic behavior. *Journal of Personality and Social Psychology*, 107(3), 393–413. <https://doi.org/10.1037/a0037207>
- Barclay, P. (2011). Competitive helping increases with the size of biological markets and invades defection. *Journal of Theoretical Biology*, 281(1), 47–55. <https://doi.org/10.1016/j.jtbi.2011.04.023>
- Barclay, P. (2013). Strategies for cooperation in biological markets, especially for humans. *Evolution and Human Behavior*, 34(3), 164–175. <https://doi.org/10.1016/j.evolhumbehav.2013.02.002>
- Barclay, P. (2016). Biological markets and the effects of partner choice on cooperation and friendship. *Current Opinion in Psychology*, 7, 33–38. <https://doi.org/10.1016/j.copsyc.2015.07.012>
- Barclay, P., & Barker, J. L. (2020). Greener than thou: People who protect the environment are more cooperative, compete to be environmental, and benefit from reputation. *Journal of Environmental Psychology*, 72, 101441. <https://doi.org/10.1016/j.jenvp.2020.101441>
- Barclay, P., & Raihani, N. J. (2016). Partner choice versus punishment in human Prisoner's Dilemmas. *Evolution and Human Behavior*, 37(4), 263–271. <https://doi.org/10.1016/j.evolhumbehav.2015.12.004>
- Barclay, P., & Willer, R. (2007). Partner choice creates competitive altruism in humans. *Proceedings of the Royal Society B: Biological Sciences*, 274(1610), 749–753. <https://doi.org/10.1098/rspb.2006.0209>
- Barker, J. L., Power, E. A., Heap, S., Puurtinen, M., & Sosis, R. (2019). Content, cost, and context: A framework for understanding human signaling systems. *Evolutionary Anthropology: Issues, News, and Reviews*, 28(2), 86–99. <https://doi.org/10.1002/evan.21768>
- Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M. T., Fitzpatrick, S., Gurven, M.,... Laurence, S. (2016). Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. *Proceedings of the National Academy of Sciences*, 113(17), 4688–4693. <https://doi.org/10.1073/pnas.1522070113>
- Bashir, N. Y., Lockwood, P., Chasteen, A. L., Nadolny, D., & Noyes, I. (2013). The ironic impact of activists: Negative stereotypes reduce social change influence. *European Journal of Social Psychology*, 43(7), 614–626. <https://doi.org/10.1002/ejsp.1983>
- Batson, C. D., Duncan, B. D., Ackerman, P., Buckley, T., & Birch, K. (1981). Is empathic emotion a source of altruistic motivation? *Journal of Personality and Social Psychology*, 40, 290–302.
- Baumard, N., André, J.-B., & Sperber, D. (2013). A mutualistic approach to morality: The evolution of fairness by partner choice. *Behavioral and Brain Sciences*, 36(1), 59–78. <https://doi.org/10.1017/S0140525X11002202>
- Bekkers, R., & Wiepking, P. (2011). A literature review of empirical studies of philanthropy: Eight mechanisms that drive charitable giving. *Nonprofit and Voluntary Sector Quarterly*, 40(5), 924–973. <https://doi.org/10.1177/0899764010380927>
- Bénabou, R., & Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review*, 96, 1652–1678.
- 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(4), 277–284. <https://doi.org/10.1016/j.evolhumbehav.2007.04.002>
- Berezkei, T., Szabo, Z. P., & Czibor, A. (2015). Abusing good intentions: Machiavellians strive for exploiting cooperators. *SAGE Open*, 5(2), 2158244015593119. <https://doi.org/10.1177/2158244015593119>
- Berge, J. W., & Arshad, F. M. (2020). *Doing well while doing good: Do firms' profit motives for doing good matter to employees?* <http://scholarspace.manoa.hawaii.edu/handle/10125/70512>
- Berman, J. Z., Barasch, A., Levine, E. E., & Small, D. A. (2018). Impediments to effective altruism: The role of subjective preferences in charitable giving. *Psychological Science*, 29(5), 834–844. <https://doi.org/10.1177/0956797617747648>
- Berman, J. Z., Levine, E. E., Barasch, A., & Small, D. A. (2015). The braggart's dilemma: On the social rewards and penalties of advertising prosocial behavior. *Journal of Marketing Research*, 52(1), 90–104. <https://doi.org/10.1509/jmr.14.0002>
- Bhogal, M. S., Farrelly, D., Galbraith, N., Manktelow, K., & Bradley, H. (2020). The role of altruistic costs in human mate choice. *Personality and Individual Differences*, 160, 109939. <https://doi.org/10.1016/j.paid.2020.109939>
- Big Charitable Gifts*. (2020, October 12). COP. <https://www.philanthropy.com/article/big-charitable-gifts>
- Bigman, Y. E., & Tamir, M. (2016). The road to heaven is paved with effort: Perceived effort amplifies moral judgment. *Journal of Experimental Psychology: General*, 145(12), 1654–1669. <https://doi.org/10.1037/xge0000230>

- Bjørkelo, B., Einarsen, S., Nielsen, M. B., & Matthiesen, S. B. (2011). Silence is golden? Characteristics and experiences of self-reported whistleblowers. *European Journal of Work and Organizational Psychology*, 20(2), 206–238. <https://doi.org/10.1080/13594320903338884>
- Blake, P. R., McAuliffe, K., Corbit, J., Callaghan, T. C., Barry, O., Bowie, A., ... Warneken, F. (2015). The ontogeny of fairness in seven societies. *Nature*, 528(7581), 258–261. <https://doi.org/10.1038/nature15703>
- Bliege Bird, R., & Power, E. A. (2015). Prosocial signaling and cooperation among Martu hunters. *Evolution and Human Behavior*, 36(5), 389–397. <https://doi.org/10.1016/j.evolhumbehav.2015.02.003>
- Bliege Bird, R., Ready, E., & Power, E. A. (2018). The social significance of subtle signals. *Nature Human Behaviour*, 2(2), 1–6. <https://doi.org/10.1038/s41562-018-0298-3>
- Bliege Bird, R., Scelza, B., Bird, D. W., & Smith, E. A. (2012). The hierarchy of virtue: Mutualism, altruism and signaling in Martu women's cooperative hunting. *Evolution and Human Behavior*, 33(1), 64–78. <https://doi.org/10.1016/j.evolhumbehav.2011.05.007>
- Bohnet, I., & Frey, B. S. (1999). Social distance and other-regarding behavior in dictator games: Comment. *American Economic Review*, 89(1), 335–339. <https://doi.org/10.1257/aer.89.1.335>
- Bolderdijk, J. W., Brouwer, C., & Cornelissen, G. (2018). When do morally motivated innovators elicit inspiration instead of irritation? *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.02362>
- Bone, J., S., A. S., & Raihani, N. J. (2014). Defectors, not norm violators, are punished by third-parties. *Biology Letters*, 10(7), 20140388. <https://doi.org/10.1098/rsbl.2014.0388>
- Boone, J. L. (1998). The evolution of magnanimity. *Human Nature*, 9(1), 21.
- Borgerhoff Mulder, M., Bowles, S., Hertz, T., Bell, A., Beise, J., ... Wiessner, P. (2009). Intergenerational wealth transmission and the dynamics of inequality in small-scale societies. *Science*, 326(5953), 682–688. <https://doi.org/10.1126/science.1178336>
- Boster, F. J., Rodriguez, J. I., Cruz, M. G., & Marshall, L. (1995). The relative effectiveness of a direct request message and a pre-giving message on friends and strangers. *Communication Research*, 22(4), 475–484. <https://doi.org/10.1177/009365095022004005>
- Bourdieu, P. (1990). *The logic of practice* (R. Nice, Trans.). Stanford University Press.
- Bowles, S., Smith, E. A., & Borgerhoff Mulder, M. (2010). The emergence and persistence of inequality in premodern societies: Introduction to the special section. *Current Anthropology*, 51(1), 7–17. <https://doi.org/10.1086/649206>
- Boyd, R. (2017). *A different kind of animal: How culture transformed our species*. Princeton University Press.
- Bradley, A., Lawrence, C., & Ferguson, E. (2018). Does observability affect prosociality? *Proceedings of the Royal Society B: Biological Sciences*, 285(1875), 20180116. <https://doi.org/10.1098/rspb.2018.0116>
- Bshary, R., & Grutter, A. S. (2006). Image scoring and cooperation in a cleaner fish mutualism. *Nature*, 441, 975–978.
- Bshary, R., & Raihani, N. J. (2017). Helping in humans and other animals: A fruitful interdisciplinary dialogue. *Proceedings of the Royal Society B: Biological Sciences*, 284(1863), 20170929. <https://doi.org/10.1098/rspb.2017.0929>
- Burtch, G., Ghose, A., & Watal, S. (2016). Secret admirers: An empirical examination of information hiding and contribution dynamics in online crowdfunding. *Information Systems Research*, 27(3), 478–496. <https://doi.org/10.1287/isre.2016.0642>
- Burum, B., Nowak, M. A., & Hoffman, M. (2020). An evolutionary explanation for ineffective altruism. *Nature Human Behaviour*, 4(12), 1245–1257. <https://doi.org/10.1038/s41562-020-00950-4>
- Camerer, C. F. (2011). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.
- Carlson, R. W., & Zaki, J. (2018). Good deeds gone bad: Lay theories of altruism and selfishness. *Journal of Experimental Social Psychology*, 75, 36–40. <https://doi.org/10.1016/j.jesp.2017.11.005>
- Caviola, L., Schubert, S., & Nemirow, J. (2020). The many obstacles to effective giving. *Judgment and Decision Making*, 15(2), 159–172.
- Chao, M. (2017). Demotivating incentives and motivation crowding out in charitable giving. *Proceedings of the National Academy of Sciences*, 114(28), 7301–7306. <https://doi.org/10.1073/pnas.1616921114>
- Chernev, A., & Blair, S. (2015). Doing well by doing good: The benevolent halo of corporate social responsibility. *Journal of Consumer Research*, 41(6), 1412–1425. <https://doi.org/10.1086/680089>
- Cowell, J. M., Lee, K., Malcolm-Smith, S., Selcuk, B., Zhou, X., & Decety, J. (2017). The development of generosity and moral cognition across five cultures. *Developmental Science*, 20(4), e12403. <https://doi.org/10.1111/desc.12403>
- Cowgill, C. M. (2017). *Good, but not a goody two-shoes: Self-presentation concerns of appearing overly moral*. Ohio University. https://etd.ohiolink.edu/apexprod/rws_olink/tr/1501/10?clear=10&p10_accession_num=ohiou1492991431186931
- Critcher, C. R., Inbar, Y., & Pizarro, D. A. (2013). How quick decisions illuminate moral character. *Social Psychological and Personality Science*, 4(3), 308–315. <https://doi.org/10.1177/1948550612457688>
- Curtin, C. M., Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M. T., Fitzpatrick, S., ... Henrich, J. (2020). Kinship intensity and the use of mental states in moral judgment across societies. *Evolution and Human Behavior*, 41(5), 415–429. <https://doi.org/10.1016/j.evolhumbehav.2020.07.002>
- De Freitas, J., DeScioli, P., Thomas, K. A., & Pinker, S. (2019). Maimonides' ladder: States of mutual knowledge and the perception of charity. *Journal of Experimental Psychology: General*, 148(1), 158–173. <https://doi.org/10.1037/xge0000507>
- Delton, A. W., & Robertson, T. E. (2016). How the mind makes welfare tradeoffs: Evolution, computation, and emotion. *Current Opinion in Psychology*, 7, 12–16. <https://doi.org/10.1016/j.copsyc.2015.06.006>

- Denis, E., Pecheux, C., & Warlop, L. (2020). When public recognition inhibits prosocial behavior: The case of charitable giving. *Nonprofit and Voluntary Sector Quarterly*, 0899764020911203. <https://doi.org/10.1177/0899764020911203>
- Effron, D. A., O'Connor, K., Leroy, H., & Lucas, B. J. (2018). From inconsistency to hypocrisy: When does 'saying one thing but doing another' invite condemnation? *Research in Organizational Behavior*, 38, 61–75. <https://doi.org/10.1016/j.riob.2018.10.003>
- Engelmann, J. M., Herrmann, E., & Tomasello, M. (2012). Five-year olds, but not chimpanzees, attempt to manage their reputations. *PLOS ONE*, 7(10), e48433. <https://doi.org/10.1371/journal.pone.0048433>
- Engelmann, J. M., Herrmann, E., & Tomasello, M. (2016). Preschoolers affect others' reputations through prosocial gossip. *British Journal of Developmental Psychology*, 34(3), 447–460. <https://doi.org/10.1111/bjdp.12143>
- Engelmann, J. M., & Rapp, D. J. (2018). The influence of reputational concerns on children's prosociality. *Current Opinion in Psychology*, 20, 92–95. <https://doi.org/10.1016/j.copsyc.2017.08.024>
- Enke, B. (2019). Kinship, cooperation, and the evolution of moral systems. *The Quarterly Journal of Economics*, 134(2), 953–1019. <https://doi.org/10.1093/qje/qjz001>
- Erlandsson, A., Wingren, M., & Andersson, P. A. (2020). Type and amount of help as predictors for impression of helpers. *PLOS ONE*, 15(12), e0243808. <https://doi.org/10.1371/journal.pone.0243808>
- Evans, A. M., & van de Calseyde, P. P. F. M. (2017). The effects of observed decision time on expectations of extremity and cooperation. *Journal of Experimental Social Psychology*, 68, 50–59. <https://doi.org/10.1016/j.jesp.2016.05.009>
- Everett, J. A. C., Faber, N. S., Savulescu, J., & Crockett, M. J. (2018). The costs of being consequentialist: Social inference from instrumental harm and impartial beneficence. *Journal of Experimental Social Psychology*, 79, 200–216. <https://doi.org/10.1016/j.jesp.2018.07.004>
- Everett, J. A. C., Pizarro, D. A., & Crockett, M. J. (2016). Inference of trustworthiness from intuitive moral judgments. *Journal of Experimental Psychology: General*, 145(6), 772–787. <https://doi.org/10.1037/xge0000165>
- Falk, A., Fehr, E., & Fischbacher, U. (2005). Driving forces behind informal sanctions. *Econometrica*, 73, 2017–2030. <https://doi.org/10.1111/j.1468-0262.2005.00644.x>
- Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, 90(4), 980–994. <https://doi.org/10.1257/aer.90.4.980>
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868), 137–140. <https://doi.org/10.1038/415137a>
- Fehrler, S., & Przepiorka, W. (2016). Choosing a partner for social exchange: Charitable giving as a signal of trustworthiness. *Journal of Economic Behavior & Organization*, 129, 157–171. <https://doi.org/10.1016/j.jebo.2016.06.006>
- Fessler, D. M. T., Barrett, H. C., Kanovsky, M., Stich, S., Holbrook, C., Henrich, J., ... Laurence, S. (2015). Moral parochialism and contextual contingency across seven societies. *Proceedings of the Royal Society B: Biological Sciences*, 282(1813), 20150907. <https://doi.org/10.1098/rspb.2015.0907>
- Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, 71(3), 397–404.
- Fisher, J. D., Nadler, A., & Whitcher-Alagna, S. (1982). Recipient reactions to aid. *Psychological Bulletin*, 91(1), 27–54. <https://doi.org/10.1037/0033-2909.91.1.27>
- Frank, R. H. (1991). *Passions within reason: The strategic role of the emotions* (new ed.). W. W. Norton.
- Fruhen, L. S., & Flin, R. (2015). Car driver attitudes, perceptions of social norms and aggressive driving behaviour towards cyclists. *Accident Analysis & Prevention*, 83, 162–170. <https://doi.org/10.1016/j.aap.2015.07.003>
- Fu, F., Hauert, C., Nowak, M. A., & Wang, L. (2008). Reputation-based partner choice promotes cooperation in social networks. *Physical Review E*, 78(2), 026117. <https://doi.org/10.1103/PhysRevE.78.026117>
- Fu, G., Heyman, G. D., Cameron, C. A., & Lee, K. (2016). Learning to be unsung heroes: Development of reputation management in two cultures. *Child Development*, 87(3), 689–699. <https://doi.org/10.1111/cdev.12494>
- Gambetta, D., & Przepiorka, W. (2014). Natural and strategic generosity as signals of trustworthiness. *PLOS ONE*, 9(5), e97533. <https://doi.org/10.1371/journal.pone.0097533>
- Genyue, F., Heyman, G. D., & Lee, K. (2011). Reasoning about modesty among adolescents and adults in China and the U.S. *Journal of Adolescence*, 34(4), 599–608. <https://doi.org/10.1016/j.adolescence.2010.10.003>
- Gino, F., Gu, J., & Zhong, C.-B. (2009). Contagion or restitution? When bad apples can motivate ethical behavior. *Journal of Experimental Social Psychology*, 45(6), 1299–1302. <https://doi.org/10.1016/j.jesp.2009.07.014>
- Gintis, H., SMITH, E. A., & Bowles, S. (2001). Costly signaling and cooperation. *Journal of Theoretical Biology*, 213(1), 103–119. <https://doi.org/10.1006/jtbi.2001.2406>
- Gneezy, A., & Epley, N. (2014). Worth keeping but not exceeding: Asymmetric consequences of breaking versus exceeding promises. *Social Psychological and Personality Science*, 5(7), 796–804. <https://doi.org/10.1177/1948550614533134>
- Gneezy, A., Gneezy, U., Riener, G., & Nelson, L. D. (2012). Pay-what-you-want, identity, and self-signaling in markets. *Proceedings of the National Academy of Sciences*, 109(19), 7236–7240. <https://doi.org/10.1073/pnas.1120893109>
- Gneezy, U., & Rustichini, A. (2000). A fine is a price. *The Journal of Legal Studies*, 29(1), 1–17. <https://doi.org/10.1086/468061>
- Guglielmo, S., & Malle, B. F. (2019). Asymmetric morality: Blame is more differentiated and more extreme than praise. *PLOS ONE*, 14(3), e0213544. <https://doi.org/10.1371/journal.pone.0213544>

- Gurven, M., Allen-Arave, W., Hill, K., & Hurtado, M. (2000). 'It's a wonderful life': Signaling generosity among the Ache of Paraguay. *Evolution and Human Behavior*, 21(4), 263–282. [https://doi.org/10.1016/S1090-5138\(00\)00032-5](https://doi.org/10.1016/S1090-5138(00)00032-5)
- Hafenbrädl, S., & Waeger, D. (2019). The business case for CSR: A trump card against hypocrisy? *Journal of Business Research*. <https://doi.org/10.1016/j.jbusres.2019.08.043>
- Halabi, S., Dovidio, J. F., & Nadler, A. (2016). Help that hurts? Perceptions of intergroup assistance. *International Journal of Intercultural Relations*, 53, 65–71. <https://doi.org/10.1016/j.ijintrel.2016.05.007>
- Hamlin, J. K. (2013). Moral judgment and action in preverbal infants and toddlers. Cdp.Sagepub.Com. <http://cdp.sagepub.com/content/22/3/186.full.pdf+html>
- Hamlin, J. K., Wynn, K., & Bloom, P. (2011). How infants and toddlers react to antisocial others. *Proceedings of the National Academy of Sciences USA*. <http://www.pnas.org/content/108/50/19931.short>
- Harbaugh, W. T. (1998). What do donations buy?: A model of philanthropy based on prestige and warm glow. *Journal of Public Economics*, 67(2), 269–284. [https://doi.org/10.1016/S0047-2727\(97\)00062-5](https://doi.org/10.1016/S0047-2727(97)00062-5)
- Harbaugh, W. T., Mayr, U., & Burghart, D. R. (2007). Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science*, 316(5831), 1622–1625. <https://doi.org/10.1126/science.1140738>
- Hardy, R. A., & Norgaard, J. R. (2016). Reputation in the Internet black market: An empirical and theoretical analysis of the Deep Web. *Journal of Institutional Economics*, 12(3), 515–539. <https://doi.org/10.1017/S1744137415000454>
- Henrich, J. (2017). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press.
- Henrich, J. (2020). *The weirdest people in the world: How the West became psychologically peculiar and particularly prosperous* (1st ed.). Allen Lane.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., ... Tracer, D. (2005). 'Economic man' in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, 28(6), 795–815. <https://doi.org/10.1017/S0140525X05000142>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, 466, 29. <https://doi.org/10.1038/466029a>
- Herrmann, E., Engelmann, J. M., & Tomasello, M. (2019). Children engage in competitive altruism. *Journal of Experimental Child Psychology*, 179, 176–189. <https://doi.org/10.1016/j.jecp.2018.11.008>
- Herrmann, E., Keupp, S., Hare, B., Vaish, A., & Tomasello, M. (2013). Direct and indirect reputation formation in nonhuman great apes (*Pan paniscus*, *Pan troglodytes*, *Gorilla gorilla*, *Pongo pygmaeus*) and human children (*Homo sapiens*). *Journal of Comparative Psychology*, 127(1), 63–75. <https://doi.org/10.1037/a0028929>
- Herrmann, B., Thoni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science*, 319(5868), 1362–1367. <https://doi.org/10.1126/science.1153808>
- Heyes, C. (2019). Précis of *Cognitive gadgets: The cultural evolution of thinking*. *Behavioral and Brain Sciences*, 42, e169. <https://doi.org/10.1017/S0140525X18002145>
- Heyman, G., Barner, D., Heumann, J., & Schenck, L. (2014). Children's sensitivity to ulterior motives when evaluating prosocial behavior. *Cognitive Science*, 38(4), 683–700. <https://doi.org/10.1111/cogs.12089>
- Heyman, G. D., Fu, G., Barner, D., Zhishan, H., Zhou, L., & Lee, K. (2016). Children's evaluation of public and private generosity and its relation to behavior: Evidence from China. *Journal of Experimental Child Psychology*, 150, 16–30. <https://doi.org/10.1016/j.jecp.2016.05.001>
- Heyman, G. D., Fu, G., & Lee, K. (2007). Evaluating claims people make about themselves: The development of skepticism. *Child Development*, 78(2), 367–375. <https://doi.org/10.1111/j.1467-8624.2007.01003.x>
- Heyman, G. D., Itakura, S., & Lee, K. (2011). Japanese and American children's reasoning about accepting credit for prosocial behavior. *Social Development*, 20(1), 171–184. <https://doi.org/10.1111/j.1467-9507.2010.00578.x>
- Hicks, C. M., Liu, D., & Heyman, G. D. (2015). Young children's beliefs about self-disclosure of performance failure and success. *British Journal of Developmental Psychology*, 33(1), 123–135. <https://doi.org/10.1111/bjdp.12077>
- Hirshleifer, J. (1987). *On the emotions as guarantors of threats and promises*. MIT Press.
- Hoffman, M., Hilbe, C., & Nowak, M. A. (2018). The signal-burying game can explain why we obscure positive traits and good deeds. *Nature Human Behaviour*, 2(6), 397–404. <https://doi.org/10.1038/s41562-018-0354-z>
- Hoffman, E., McCabe, K., Shachat, K., & Smith, V. (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior*, 7, 346–380.
- Hoffman, M., Yoeli, E., & Nowak, M. A. (2015). Cooperate without looking: Why we care what people think and not just what they do. *Proceedings of the National Academy of Sciences*, 112(6), 1727–1732. <https://doi.org/10.1073/pnas.1417904112>
- House, B. R., Kanngiesser, P., Barrett, H. C., Broesch, T., Cebioglu, S., Crittenden, A. N., ... Silk, J. B. (2020a). Universal norm psychology leads to societal diversity in prosocial behaviour and development. *Nature Human Behaviour*, 4(1), 36–44. <https://doi.org/10.1038/s41562-019-0734-z>
- House, B. R., Kanngiesser, P., Barrett, H. C., Yilmaz, S., Smith, A. M., Sebastian-Enesco, C., ... Silk, J. B. (2020b). Social norms and cultural diversity in the development of third-party punishment. *Proceedings of the Royal Society B: Biological Sciences*, 287(1925), 20192794. <https://doi.org/10.1098/rspb.2019.2794>

- House, B. R., Silk, J. B., Henrich, J., Barrett, H. C., Scelza, B. A., Boyette, A. H., ... Laurence, S. (2013). Ontogeny of prosocial behavior across diverse societies. *Proceedings of the National Academy of Sciences*, 110(36), 14586–14591. <https://doi.org/10.1073/pnas.1221217110>
- Howe, L. C., & Monin, B. (2017). Healthier than thou? ‘Practicing what you preach’ backfires by increasing anticipated devaluation. *Journal of Personality and Social Psychology*, 112(5), 718–735. <https://doi.org/10.1037/pspi0000085>
- Imada, H. (2020). Preference for anonymous giving. *Letters on Evolutionary Behavioral Science*, 11(1), 22–26. <https://doi.org/10.5178/lebs.2020.76>
- Irwin, K., & Horne, C. (2013). A normative explanation of antisocial punishment. *Social Science Research*, 42(2), 562–570. <https://doi.org/10.1016/j.ssresearch.2012.10.004>
- Jackson, M. O., Rodriguez-Barraquer, T., & Tan, X. (2012). Social capital and social quilts: Network patterns of favor exchange. *American Economic Review*, 102(5), 1857–1897. <https://doi.org/10.1257/aer.102.5.1857>
- Jenni, K., & Loewenstein, G. (1997). Explaining the identifiable victim effect. *Journal of Risk and Uncertainty*, 14(3), 235–257. <https://doi.org/10.1023/A:1007740225484>
- Johnsen, Å. A., & Kvaløy, O. (2016). Does strategic kindness crowd out prosocial behavior? *Journal of Economic Behavior & Organization*, 132, 1–11. <https://doi.org/10.1016/j.jebo.2016.09.016>
- Johnson, S. (2018). *Dimensions of altruism: Do evaluations of prosocial behavior track social good or personal sacrifice?* (SSRN Scholarly Paper ID 3277444). Social Science Research Network. <https://doi.org/10.2139/ssrn.3277444>
- Johnson, T., Dawes, C. T., Fowler, J. H., McElreath, R., & Smirnov, O. (2009). The role of egalitarian motives in altruistic punishment. *Economics Letters*, 102(3), 192–194. <https://doi.org/10.1016/j.econlet.2009.01.003>
- Johnson, S. G. B., & Park, S. Y. (2019). *Moral signaling through donations of money and time*. PsyArXiv. <https://doi.org/10.31234/osf.io/tg9xs>
- Jones, D., & Linardi, S. (2014). Wallflowers: Experimental evidence of an aversion to standing out. *Management Science*, 60(7), 1757–1771.
- Jordan, J. J., Hoffman, M., Nowak, M. A., & Rand, D. G. (2016). Uncalculating cooperation is used to signal trustworthiness. *Proceedings of the National Academy of Sciences*, 113(31), 8658–8663. <https://doi.org/10.1073/pnas.1601280113>
- Jordan, J. J., Rand, D. G., Arbesman, S., Fowler, J. H., & Christakis, N. A. (2013). Contagion of cooperation in static and fluid social networks. *PLOS ONE*, 8(6), e66199. <https://doi.org/10.1371/journal.pone.0066199>
- Jordan, J. J., Sommers, R., Bloom, P., & Rand, D. G. (2017). Why do we hate hypocrites? Evidence for a theory of false signaling. *Psychological Science*, 28(3), 356–368. <https://doi.org/10.1177/0956797616685771>
- Kanagawa, C., Cross, S. E., & Markus, H. R. (2001). ‘Who am I?’ The cultural psychology of the conceptual self. *Personality and Social Psychology Bulletin*, 27(1), 90–103. <https://doi.org/10.1177/0146167201271008>
- Kandori, M. (1992). Social norms and community enforcement. *The Review of Economic Studies*, 59(1), 63–80. <https://doi.org/10.2307/2297925?ref=search-gateway:3a8c4b0a46e82c3f16920b6ecbabe03>
- Kawamura, Y., & Kusumi, T. (2020). Altruism does not always lead to a good reputation: A normative explanation. *Journal of Experimental Social Psychology*, 90, 104021. <https://doi.org/10.1016/j.jesp.2020.104021>
- Kawamura, Y., Ohtsubo, Y., & Kusumi, T. (2020). Effects of cost and benefit of prosocial behavior on reputation. *Social Psychological and Personality Science*, 1948550620929163. <https://doi.org/10.1177/1948550620929163>
- Kim, T., Zhang, T., & Norton, M. I. (2019). Pettiness in social exchange. *Journal of Experimental Psychology: General*, 148(2), 361–373. <https://doi.org/10.1037/xge0000463>
- Kirgios, E. L., Chang, E. H., Levine, E. E., Milkman, K. L., & Kessler, J. B. (2020). Forgoing earned incentives to signal pure motives. *Proceedings of the National Academy of Sciences*. <https://doi.org/10.1073/pnas.2000065117>
- Klein, N., & Epley, N. (2014). The topography of generosity: Asymmetric evaluations of prosocial actions. *Journal of Experimental Psychology: General*, 143(6), 2366–2379. <https://doi.org/10.1037/xge0000025>
- Klein, N., Grossmann, I., Uskul, A. K., Kraus, A. A., & Epley, N. (2015). It pays to be nice, but not really nice: Asymmetric reputations from prosociality across 7 countries. *Judgment and Decision Making*, 10(4), 10.
- Krems, J. A., Williams, K. E., Aktipis, A., & Kenrick, D. T. (2021). Friendship jealousy: One tool for maintaining friendships in the face of third-party threats?. *Journal of Personality and Social Psychology*, 120(4), 977.
- Krull, D. S., Seger, C. R., & Silvera, D. H. (2008). Smile when you say that: Effects of willingness on dispositional inferences. *Journal of Experimental Social Psychology*, 44(3), 735–742. <https://doi.org/10.1016/j.jesp.2007.05.004>
- Kurz, T., Prosser, A. M. B., Rabinovich, A., & O’Neill, S. (2020). Could vegans and lycra cyclists be bad for the planet? Theorizing the role of moralized minority practice identities in processes of societal-level change. *Journal of Social Issues*, 76(1), 86–100. <https://doi.org/10.1111/josi.12366>
- Kutlaca, M., Becker, J., & Radke, H. (2020). A hero for the outgroup, a black sheep for the ingroup: Societal perceptions of those who confront discrimination. *Journal of Experimental Social Psychology*, 88, 103832. <https://doi.org/10.1016/j.jesp.2019.103832>
- Lamba, S., & Mace, R. (2011). Demography and ecology drive variation in cooperation across human populations. *Proceedings of the National Academy of Sciences*, 108(35), 14426–14430. <https://doi.org/10.1073/pnas.1105186108>
- Law, K. F., Campbell, D., & Gaesser, B. (2019). *Biased benevolence: The morality of effective altruism* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/qzx67>

- Lee, R. B. (1969). *Eating Christmas in the Kalahari*. 4.
- Leimgruber, K. L., Shaw, A., Santos, L. R., & Olson, K. R. (2012). Young children are more generous when others are aware of their actions. *PLOS ONE*, 7(10), e48292. <https://doi.org/10.1371/journal.pone.0048292>
- Levine, E. E., Barasch, A., Rand, D., Berman, J. Z., & Small, D. A. (2018). Signaling emotion and reason in cooperation. *Journal of Experimental Psychology: General*, 147(5), 702–719. <https://doi.org/10.1037/xge0000399>
- Levy, N. (2020). Virtue signalling is virtuous. *Synthese*. <https://doi.org/10.1007/s11229-020-02653-9>
- Li, Y. (2014). The evolution of reputation-based partner-switching behaviors with a cost. *Scientific Reports*, 4(1), 1–10. <https://doi.org/10.1038/srep05957>
- Lin-Healy, F., & Small, D. A. (2012). Cheapened altruism: Discounting personally affected prosocial actors. *Organizational Behavior and Human Decision Processes*, 117(2), 269–274. <https://doi.org/10.1016/j.obhdp.2011.11.006>
- Luhrmann, T. M. (2011). Toward an anthropological theory of mind. *Suomen Antropologi: Journal of the Finnish Anthropological Society*, 36(4), 5–69.
- Lyle, H. F., & Smith, E. A. (2014). The reputational and social network benefits of prosociality in an Andean community. *Proceedings of the National Academy of Sciences*, 111(13), 4820–4825. <https://doi.org/10.1073/pnas.1318372111>
- Macfarlan, S. J., & Lyle, H. F. (2015). Multiple reputation domains and cooperative behaviour in two Latin American communities. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1683), 20150009. <https://doi.org/10.1098/rstb.2015.0009>
- Macfarlan, S. J., Quinlan, R., & Remiker, M. (2013). Cooperative behaviour and prosocial reputation dynamics in a Dominican village. *Proceedings of the Royal Society B: Biological Sciences*, 280(1761), 20130557. <https://doi.org/10.1098/rspb.2013.0557>
- Macfarlan, S. J., Remiker, M., & Quinlan, R. (2012). Competitive altruism explains labor exchange variation in a Dominican community. *Current Anthropology*, 53(1), 118–124. <https://doi.org/10.1086/663700>
- Makov, T., & Newman, G. E. (2016). Economic gains stimulate negative evaluations of corporate sustainability initiatives. *Nature Climate Change*, 6(9), 844–846. <https://doi.org/10.1038/nclimate3033>
- Mathew, S., & Boyd, R. (2014). The cost of cowardice: Punitive sentiments towards free riders in Turkana raids. *Evolution and Human Behavior*, 35(1), 58–64. <https://doi.org/10.1016/j.evolhumbehav.2013.10.001>
- McAndrew, F. T., & Perilloux, C. (2012). Is self-sacrificial competitive altruism primarily a male activity? *Evolutionary Psychology*, 10(1), 147470491201000100. <https://doi.org/10.1177/147470491201000107>
- McAuliffe, K., Wrangham, R., Glowacki, L., & Russell, A. F. (2015). When cooperation begets cooperation: The role of key individuals in galvanizing support. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1683), 20150012. <https://doi.org/10.1098/rstb.2015.0012>
- McManus, R. M., Kleiman-Weiner, M., & Young, L. (2020). What we owe to family: The impact of special obligations on moral judgment. *Psychological Science*, 31(3), 227–242. <https://doi.org/10.1177/0956797619900321>
- McNamara, J. M., & Barta, Z. (2020). Behavioural flexibility and reputation formation. *Proceedings of the Royal Society B: Biological Sciences*, 287(1939), 20201758. <https://doi.org/10.1098/rspb.2020.1758>
- McNamara, J. M., Barta, Z., Fromhage, L., & Houston, A. I. (2008). The coevolution of choosiness and cooperation. *Nature*, 451(7175), 189–192. <https://doi.org/10.1038/nature06455>
- Melis, A. P., & Semmann, D. (2010). How is human cooperation different? *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365(1553), 2663–2674. <https://doi.org/10.1098/rstb.2010.0157>
- Mellström, C., & Johannesson, M. (2008). Crowding out in blood donation: Was Titmuss right? *Journal of the European Economic Association*, 6(4), 845–863. <https://doi.org/10.1162/JEEA.2008.6.4.845>
- Melnikoff, D. E., & Bailey, A. H. (2018). Preferences for moral vs. Immoral traits in others are conditional. *Proceedings of the National Academy of Sciences*, 115(4), E592–E600. <https://doi.org/10.1073/pnas.1714945115>
- Milinski, M., Semmann, D., & Krambeck, H. (2002a). Donors to charity gain in both indirect reciprocity and political reputation. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269(1494), 881.
- Milinski, M., Semmann, D., & Krambeck, H. (2002b). Reputation helps solve the ‘tragedy of the commons’. *Nature*, 415(6870), 424–426.
- Mills, C. M., & Keil, F. C. (2005). The development of cynicism. *Psychological Science*, 16(5), 385–390. <https://doi.org/10.1111/j.0956-7976.2005.01545.x>
- Minson, J. A., & Monin, B. (2012). Do-gooder derogation: Disparaging morally motivated minorities to defuse anticipated reproach. *Social Psychological and Personality Science*, 3(2), 200–207. <https://doi.org/10.1177/1948550611415695>
- Mokos, J., & Scheuring, I. (2019). Altruism, costly signaling, and withholding information in a sport charity campaign. *Evolution, Mind and Behaviour*, 17(1), 10–18. <https://doi.org/10.1556/2050.2019.00007>
- Molho, C., Tybur, J. M., Van Lange, P. A. M., & Balliet, D. (2020). Direct and indirect punishment of norm violations in daily life. *Nature Communications*, 11(1), 3432. <https://doi.org/10.1038/s41467-020-17286-2>
- Monin, B. (2007). Holier than me? Threatening Social Comparison in the Moral Domain. *Revue internationale de psychologie sociale*, Tome 20(1), 53–68.
- Monin, B., Sawyer, P. J., & Marquez, M. J. (2008). The rejection of moral rebels: Resenting those who do the right thing. *Journal of Personality and Social Psychology*, 95(1), 76–93. <https://doi.org/10.1037/0022-3514.95.1.76>

- Monteleagre, A., Bush, L., Moss, D., Pizarro, D., & Jimenez-Leal, W. (2020). *Does maximizing good make people look bad?* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/2zbax>
- Moussaoui, L. S., Herren-Vichet, J., Baldelli, S., Waldvogel, S., & Desrichard, O. (2019). Examining if sending text messages to thank blood donors for saving lives is a 'great idea'? Evidence from a randomized field experiment. *Journal of Nonprofit & Public Sector Marketing*, 31(5), 486–506. <https://doi.org/10.1080/10495142.2018.1526753>
- Nadler, A., & Halabi, S. (2006). Intergroup helping as status relations: Effects of status stability, identification, and type of help on receptivity to high-status group's help. *Journal of Personality and Social Psychology*, 91(1), 97–110. <https://doi.org/10.1037/0022-3514.91.1.97>
- Nelissen, R. M. A. (2008). The price you pay: Cost-dependent reputation effects of altruistic punishment. *Evolution and Human Behavior*, 29(4), 242–248. <https://doi.org/10.1016/j.evolhumbehav.2008.01.001>
- Newman, G. E., & Cain, D. M. (2014). Tainted altruism. *Psychological Science*, 25, 648.
- Newman, G. E., & Jeremy Shen, Y. (2012). The counterintuitive effects of thank-you gifts on charitable giving. *Journal of Economic Psychology*, 33(5), 973–983. <https://doi.org/10.1016/j.joep.2012.05.002>
- Nikiforakis, N., & Engelmann, D. (2011). Altruistic punishment and the threat of feuds. *Journal of Economic Behavior & Organization*, 78(3), 319–332. <https://doi.org/10.1016/j.jebo.2011.01.017>
- Ohtsuki, H., & Iwasa, Y. (2006). The leading eight: Social norms that can maintain cooperation by indirect reciprocity. *Journal of Theoretical Biology*, 239(4), 435–444. <https://doi.org/10.1016/j.jtbi.2005.08.008>
- Olivola, C. Y., & Shafir, E. (2013). The martyrdom effect: When pain and effort increase prosocial contributions. *Journal of Behavioral Decision Making*, 26(1), 91–105. <https://doi.org/10.1002/bdm.767>
- Ostrom, E., Burger, J., Field, C., Norgaard, R., & Policansky, D. (1999). Revisiting the commons: Local lessons, global challenges. *Science (New York)*, 284(5412), 278.
- Parks, C. D., & Stone, A. B. (2010). The desire to expel unselfish members from the group. *Journal of Personality and Social Psychology*, 99(2), 303–310. <https://doi.org/10.1037/a0018403>
- Peacey, M. W., & Sanders, M. (2014). *Masked heroes: Endogenous anonymity in charitable giving*. University of Bristol: Centre for Market and Public Organisation.
- Penn, D. J. (2003). The evolutionary roots of our environmental problems: Toward a Darwinian ecology. *The Quarterly Review of Biology*, 78(3), 275–301. <https://doi.org/10.1086/377051>
- Pfeiffer, T., Rutte, C., Killingback, T., Taborsky, M., & Bonhoeffer, S. (2005). Evolution of cooperation by generalized reciprocity. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 272(1568), 1115–1120. <https://doi.org/10.1098/rspb.2004.2988>
- Piazza, J., & Bering, J. (2008). Concerns about reputation via gossip promote generous allocations in an economic game. *Evolution and Human Behavior*, 29(3), 172–178.
- Pizarro, D., Uhlmann, E., & Salovey, P. (2003). Asymmetry in judgments of moral blame and praise: The role of perceived metadesires. *Psychological Science*, 14(3), 267–272. <https://doi.org/10.1111/1467-9280.03433>
- Pleasant, A., & Barclay, P. (2018). Why hate the good guy? Antisocial punishment of high cooperators is greater when people compete to be chosen. *Psychological Science*, 29(6), 868–876. <https://doi.org/10.1177/0956797617752642>
- Power, E. A., & Ready, E. (2018). Building bigness: Reputation, prominence, and social capital in rural South India. *American Anthropologist*, 120(3), 444–459. <https://doi.org/10.1111/aman.13100>
- Przepiorka, W., Norbutas, L., & Corten, R. (2017). Order without law: Reputation promotes cooperation in a cryptomarket for illegal drugs. *European Sociological Review*, 33(6), 752–764. <https://doi.org/10.1093/esr/jcx072>
- Quillien, T. (2019). Universal modesty in signal-burying games. *Proceedings of the Royal Society B: Biological Sciences*, 286(1906), 20190985. <https://doi.org/10.1098/rspb.2019.0985>
- Raihani, N. J. (2014). Hidden altruism in a real-world setting. *Biology Letters*, 10, 20130884.
- Raihani, N. J., & Aitken, D. (2011). Uncertainty, rationality and cooperation in the context of climate change. *Climatic Change*, 108(1–2), 47–55. <https://doi.org/10.1007/s10584-010-0014-4>
- Raihani, N. J., & Barclay, P. (2016). Exploring the trade-off between quality and fairness in human partner choice. *Royal Society Open Science*, 3(11), 160510. <https://doi.org/10.1098/rsos.160510>
- Raihani, N. J., & Bshary, R. (2015). Why humans might help strangers. *Frontiers in Behavioral Neuroscience*, 9, 2531. <https://doi.org/10.3389/fnbeh.2015.00039>
- Raihani, N. J., & Bshary, R. (2019). Punishment: One tool, many uses. *Evolutionary Human Sciences*, 1. <https://doi.org/10.1017/ehs.2019.12>
- Raihani, N. J., & Hart, T. (2010). Free-riders promote free-riding in a real-world setting. *Oikos*. <http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0706.2010.18279.x/pdf>
- Raihani, N. J., & Smith, S. (2015). Competitive helping in online giving. *Current Biology*, 25(9), 1183–1186. <https://doi.org/10.1016/j.cub.2015.02.042>
- Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in Cognitive Sciences*, 17, 413–425.
- Reed, L. I., Zeglen, K. N., & Schmidt, K. L. (2012). Facial expressions as honest signals of cooperative intent in a one-shot anonymous Prisoner's Dilemma game. *Evolution and Human Behavior*, 33(3), 200–209. <https://doi.org/10.1016/j.evolhumbehav.2011.09.003>

- Resnick, P., Zeckhauser, R., Swanson, J., & Lockwood, K. (2006). The value of reputation on eBay: A controlled experiment. *Experimental Economics*, 9(2), 79–101. <https://doi.org/10.1007/s10683-006-4309-2>
- Roberts, G. (1998). Competitive altruism: From reciprocity to the handicap principle. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1394), 427–431. <https://doi.org/10.1098/rspb.1998.0312>
- Roberts, G. (2005). Cooperation through interdependence. *Animal Behaviour*, 70(4), 901–908.
- Roberts, G. (2015). Partner choice drives the evolution of cooperation via indirect reciprocity. *PLoS ONE*, 10(6), e0129442. <https://doi.org/10.1371/journal.pone.0129442>
- Roberts, G. (2020). Honest signaling of cooperative intentions. *Behavioral Ecology*. <https://doi.org/10.1093/beheco/araa035>
- Salvadori, E., Blazsekova, T., Volein, A., Karap, Z., Tatone, D., Mascaro, O., & Csibra, G. (2015). Probing the strength of infants' preference for helpers over hinderers: Two replication attempts of Hamlin and Wynn (2011). *PLoS ONE*, 10(11), e0140570. <https://doi.org/10.1371/journal.pone.0140570>
- Samu, F., Számádó, S., & Takács, K. (2020). Scarce and directly beneficial reputations support cooperation. *Scientific Reports*, 10(1), 11486. <https://doi.org/10.1038/s41598-020-68123-x>
- Savary, J., & Goldsmith, K. (2020). Unobserved altruism: How self-signaling motivations and social benefits shape willingness to donate. *Journal of Experimental Psychology: Applied*, 26(3), 538–550. <https://doi.org/10.1037/xap0000261>
- Savary, J., Li, C. X., & Newman, G. E. (2020). Exalted purchases or tainted donations? Self-signaling and the evaluation of charitable incentives. *Journal of Consumer Psychology*, 30(4), 671–679. <https://doi.org/10.1002/jcpsy.1157>
- Schlingloff, L., Csibra, G., & Tatone, D. (2020). Do 15-month-old infants prefer helpers? A replication of Hamlin et al. (2007). *Royal Society Open Science*, 7(4), 191795. <https://doi.org/10.1098/rsos.191795>
- Schulz, J. F., Bahrami-Rad, D., Beauchamp, J. P., & Henrich, J. (2019). The Church, intensive kinship, and global psychological variation. *Science*, 366(6466). <https://doi.org/10.1126/science.aau5141>
- Semmann, D., Krambeck, H., & Milinski, M. (2004). Strategic investment in reputation. *Behavioral Ecology and Sociobiology*, 56(3), 248–252.
- Sen, S., Du, S., & Bhattacharya, C. (2016). Corporate social responsibility: A consumer psychology perspective. *Current Opinion in Psychology*, 10, 70–75. <https://doi.org/10.1016/j.copsyc.2015.12.014>
- Shackelford, T. K., & Buss, D. M. (1996). Betrayal in mateships, friendships, and coalitions. *Personality & Social Psychology Bulletin*, 22, 1151–1164.
- Shaw, A., & Olson, K. (2015). Whose idea is it anyway? The importance of reputation in acknowledgement. *Developmental Science*, 18(3), 502–509. <https://doi.org/10.1111/desc.12234>
- Shi, L. (2011). Monetary rewards, image concern, and intrinsic motivation: Evidence from a survey on blood donation. Working Paper, 26.
- Siegel, J. Z., Crockett, M. J., & Dolan, R. J. (2017). Inferences about moral character moderate the impact of consequences on blame and praise. *Cognition*, 167, 201–211. <https://doi.org/10.1016/j.cognition.2017.05.004>
- Siem, B., & Stürmer, S. (2018). Attribution of egoistic versus altruistic motives to acts of helping. *Social Psychology*, 50(1), 53–66. <https://doi.org/10.1027/1864-9335/a000360>
- Silk, J. B. (2003). Cooperation without counting: The puzzle of friendship. In P. Hammerstein (Ed.), *Genetic and cultural evolution of cooperation* (pp. 37–54). MIT Press.
- Silver, I., Newman, G., & Small, D. A. (2021). Inauthenticity aversion: Moral reactance toward tainted actors, actions, and objects. *Consumer Psychology Review*, 4(1), 70–82. <https://doi.org/10.1002/arcp.1064>
- Silver, I. M., & Shaw, A. (2018). Pint-sized public relations: The development of reputation management. *Trends in Cognitive Sciences*, 22(4), 277–279. <https://doi.org/10.1016/j.tics.2018.01.006>
- Simpson, B., White, K., & Laran, J. (2018). When public recognition for charitable giving backfires: The role of independent self-construal. *Journal of Consumer Research*, 44(6), 1257–1273. <https://doi.org/10.1093/jcr/ucx101>
- Singh, M., & Hoffman, M. (2021). *Commitment and impersonation: A reputation-based theory of principled behavior*. PsyArXiv. <https://doi.org/10.31234/osf.io/ua57r>
- Small, D. A., & Cryder, C. (2016). Prosocial consumer behavior. *Current Opinion in Psychology*, 10, 107–111. <https://doi.org/10.1016/j.copsyc.2016.01.001>
- Smith, E. A., & Bird, R. L. B. (2000). Turtle hunting and tombstone opening: Public generosity as costly signaling. *Evolution and Human Behavior*, 21(4), 245–261. [https://doi.org/10.1016/S1090-5138\(00\)00031-3](https://doi.org/10.1016/S1090-5138(00)00031-3)
- Smith, E. A., Bird, R. B., & Bird, D. W. (2003). The benefits of costly signaling: Meriam turtle hunters. *Behavioral Ecology*, 14(1), 116–126. <https://doi.org/10.1093/beheco/14.1.116>
- Sparkman, G., & Attari, S. Z. (2020a). Credibility, communication, and climate change: How lifestyle inconsistency and do-gooder derogation impact decarbonization advocacy. *Energy Research & Social Science*, 59, 101290. <https://doi.org/10.1016/j.erss.2019.101290>
- Sparkman, G., & Attari, S. Z. (2020b). Credibility, communication, and climate change: How lifestyle inconsistency and do-gooder derogation impact decarbonization advocacy. *Energy Research & Social Science*, 59, 101290. <https://doi.org/10.1016/j.erss.2019.101290>
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, 25(4), 359–393. <https://doi.org/10.1111/j.1468-0017.2010.01394.x>

- Srna, S., Barasch, A., & Small, D. (2017). When modesty pays. *ACR North American Advances*, NA-45. <https://www.acrwebsite.org/volumes/1024799/volumes/v45/NA-45>
- Stouten, J., van Dijke, M., Mayer, D. M., De Cremer, D., & Euwema, M. C. (2013). Can a leader be seen as too ethical? The curvilinear effects of ethical leadership. *The Leadership Quarterly*, 24(5), 680–695. <https://doi.org/10.1016/j.leaqua.2013.05.002>
- Sylwester, K., Herrmann, B., & Bryson, J. (2013). Homo Homini Lupus? Explaining antisocial punishment. *Journal of Neuroscience, Psychology, and Economics*, 6, 167–188.
- Sylwester, K., & Roberts, G. (2010). Cooperators benefit through reputation-based partner choice in economic games. *Biology Letters*, 6(5), 659–662. <https://doi.org/10.1098/rsbl.2010.0209>
- Sylwester, K., & Roberts, G. (2013). Reputation-based partner choice is an effective alternative to indirect reciprocity in solving social dilemmas. *Evolution and Human Behavior*, 34, 201–206.
- Tasimi, A., Dominguez, A., & Wynn, K. (2015). Do-gooder derogation in children: The social costs of generosity. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01036>
- Tognetti, A., Berticat, C., Raymond, M., & Faurie, C. (2012). Sexual selection of human cooperative behaviour: An experimental study in rural Senegal. *PLoS ONE*, 7, e44403.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(5), 675–691. <https://doi.org/10.1017/S0140525X05000129>
- Torelli, C. J., Monga, A. B., & Kaikati, A. M. (2012). Doing poorly by doing good: Corporate social responsibility and brand concepts. *Journal of Consumer Research*, 38(5), 948–963. <https://doi.org/10.1086/660851>
- Van Vugt, M. (2009). Averting the tragedy of the commons: Using social psychological science to protect the environment. *Current Directions in Psychological Science*, 18(3), 169–173. <https://doi.org/10.1111/j.1467-8721.2009.01630.x>
- Van Vugt, M., & Hardy, C. L. (2009). Cooperation for reputation: Wasteful contributions as costly signals in public goods. *Group Processes Intergroup Relations*. <http://gpi.sagepub.com/content/early/2009/11/13/1368430209342258.full.pdf>
- Visserman, M. L., Righetti, F., Impett, E. A., Keltner, D., & Van Lange, P. A. M. (2018). It's the motive that counts: Perceived sacrifice motives and gratitude in romantic relationships. *Emotion*, 18(5), 625–637. <https://doi.org/10.1037/emo0000344>
- von Rueden, C. R., & Jaeggi, A. V. (2016). Men's status and reproductive success in 33 nonindustrial societies: Effects of subsistence, marriage system, and reproductive strategy. *Proceedings of the National Academy of Sciences*, 113(39), 10824–10829. <https://doi.org/10.1073/pnas.1606800113>
- von Rueden, C. R., Redhead, D., O'Gorman, R., Kaplan, H., & Gurven, M. (2019). The dynamics of men's cooperation and social status in a small-scale society. *Proceedings of the Royal Society B: Biological Sciences*, 286(1908), 20191367. <https://doi.org/10.1098/rspb.2019.1367>
- Wang, X., & Tong, L. (2015). Hide the light or let it shine? Examining the factors influencing the effect of publicizing donations on donors' happiness. *International Journal of Research in Marketing*, 32(4), 418–424. <https://doi.org/10.1016/j.ijresmar.2015.06.001>
- Warneken, F., Sebastián-Enesco, C., Benjamin, N. E., & Pieloch, K. A. (2019). Pay to play: Children's emerging ability to use acts of generosity for selfish ends. *Journal of Experimental Child Psychology*, 188, 104675. <https://doi.org/10.1016/j.jecp.2019.104675>
- Weiss, A., Michels, C., Burgmer, P., Mussweiler, T., Ockenfels, A., & Hofmann, W. (2020). Trust in everyday life. *Journal of Personality and Social Psychology*. <https://doi.org/10.1037/pspi0000334>
- Yamagishi, T., Hashimoto, H., Cook, K. S., Kiyonari, T., Shinada, M., Mifune, N., ... Li, Y. (2012). Modesty in self-presentation: A comparison between the USA and Japan: Modesty in self-presentation. *Asian Journal of Social Psychology*, 15(1), 60–68. <https://doi.org/10.1111/j.1467-839X.2011.01362.x>
- Yoeli, E., Hoffman, M., Rand, D. G., & Nowak, M. A. (2013). Powering up with indirect reciprocity in a large-scale field experiment. *Proceedings of the National Academy of Sciences*, 110(Supplement 2), 10424–10429. <https://doi.org/10.1073/pnas.1301210110>
- Yoon, Y., Gürhan-Canli, Z., & Schwarz, N. (2006). The effect of corporate social responsibility (CSR) activities on companies with bad reputations. *Journal of Consumer Psychology*, 16(4), 377–390. https://doi.org/10.1207/s15327663jcp1604_9
- Yuan, M., Wu, J., & Kou, Y. (2018). Donors' social class and their prosocial reputation. *Social Psychology*, 49(4), 205–218. <https://doi.org/10.1027/1864-9335/a000342>
- Yudkin, D. A., Prosser, A. M. B., & Crockett, M. J. (2019). Actions speak louder than outcomes in judgments of prosocial behavior. *Emotion*, 19(7), 1138–1147. <https://doi.org/10.1037/emo0000514>
- Zane, D. M., Irwin, J. R., & Reczek, R. W. (2016). Do less ethical consumers denigrate more ethical consumers? The effect of willful ignorance on judgments of others. *Journal of Consumer Psychology*, 26(3), 337–349. <https://doi.org/10.1016/j.jcps.2015.10.002>