Journal of Environmental Psychology 72 (2020) 101441

Contents lists available at ScienceDirect

Journal of Environmental Psychology

journal homepage: www.elsevier.com/locate/jep

Greener Than Thou: People who protect the environment are more cooperative, compete to be environmental, and benefit from reputation

Pat Barclay [a,*], Jessica L. Barker [b,c,d]

[a] Department of Psychology, University of Guelph, 50 Stone Rd. E. Guelph, ON, N1G 2W1, Canada
[b] The Behavioural Insights Team, 4 Matthew Parker St., London, SW1H 9NP, UK
[c] Interacting Minds Centre, Aarhus University, Jens Chr. Skous Vej 4, Aarhus, Denmark
[d] Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, 85721-0088, USA

ARTICLE INFO

Handling Editor: Sander van der Linden

Keywords:
Reputation
Cooperation
Biological markets
Partner choice
Competitive altruism
Evolutionary psychology
Costly signaling
Prosocial behaviour
Sustainability
Environmentalism

ABSTRACT

Protecting the environment is a social dilemma: environmental protection benefits everyone but is individually costly. We propose that protecting the environment is similar to other types of cooperation, in that environmentalism functions as a signal of one's willingness to cooperate with others. We test several novel predictions from this hypothesis. We used a mathematical model to show that environmentalism can indicate one's valuation of others and thus one's cooperative intent. We found support for this prediction in two online studies, and then conducted two laboratory studies to extend the idea that environmentalism signals one's willingness to cooperate. Participants donated more to an environmental charity when donations were public than when anonymous, but they donated the most when competing to be chosen by an observer for a subsequent cooperative game. In other words, people competed to donate more to the environment. Bigger donors benefited, as they were subsequently chosen more often and received more cooperation from their partners. Partners benefited from choosing environmental donors: bigger donors cooperated more with subsequent partners, such that environmental donations were reliably informative about participants' future cooperativeness. We compare multiple theories about why people behave environmentally (indirect reciprocity, signal of wealth, signal of cooperative intent), and find most support for our proposed theory of signaling cooperative intent. By understanding the function of environmental behaviour and stimulating competitive giving, we can increase people's support for environmental and other charitable causes.

1. Introduction

Protecting the environment is a social dilemma (Griskevicius, Tybur, & Van den Bergh, 2010; Milinski, Semmann, Krambeck, & Marotzke, 2006). Whether they recognize it or not, everyone benefits from clean air, clean water, and abundant natural resources. However, environmentalism is individually costly because people must restrain themselves from polluting, littering and over-using limited resources. Such restraint requires long-term thinking, seeing the broader context of one's actions, and concern for the effects upon others. Researchers have identified many psychological factors involved in environmentalism, including values, norms, goals, attitudes, "warm glow", and biospheric concern (e.g., de Groot & Steg, 2008; Gifford & Nilsson, 2014; Steg, Bolderdijk, Keizer, & Perlaviciute, 2014; Steg & Vlek, 2009; van der Linden, 2018). However, to foster these proximate psychological factors, and thus to foster greater protection of the environment, we must understand their ultimate function (Barclay, 2012a): that is, the reasons why this proximate psychology exists at all (for the distinction between proximate mechanisms and ultimate function, see Scott-Phillips, Dickins, & West, 2011; Tinbergen, 1968; Section 1.3). What is the ultimate function of possessing a pro-environmental psychology (i.e., a psychology that, for whatever proximate reason, values the

* Corresponding author.
E-mail addresses: barclayp@uoguelph.ca (P. Barclay), jessiebarker@gmail.com (J.L. Barker).
URL: https://www.patbarclay.com (P. Barclay).

[1] The term "reputation" can be defined simply as the opinion that others have about oneself based on past behaviour (e.g., having a "good reputation"). It is also often used as shorthand to refer to (1) "opportunities to earn a reputation" (i.e. to earn good opinions from others), henceforth "reputational opportunities"; and (2) "the benefits one receives from earning a good reputation" (i.e. benefits from others having a good opinion of oneself), henceforth "reputational benefits". In this paper, we do not imply the proximate psychological experience of "concern for one's reputation" (i.e., "a concern about other's opinions about oneself"), which is alleged to motivate behaviour; see Section 1.3 for details. For more on these distinctions, see Barclay (2020). We invite readers who use these terms differently to substitute their preferred terms.

https://doi.org/10.1016/j.jenvp.2020.101441
Received 11 May 2019; Received in revised form 20 May 2020; Accepted 23 May 2020
Available online 27 May 2020

environment)? What payoffs result in it persisting despite the cost? A promising candidate is the benefits of a good reputation.[1]

Reputational opportunities are important in promoting cooperative behaviour (e.g., Andreoni & Bernheim, 2009; Barclay & Willer, 2007; Fehrler & Przepiorka, 2013; Feinberg, Willer, & Schultz, 2014; Milinski, Semmann, & Krambeck, 2002; Wu, Balliet, & Van Lange, 2016). People are more generous when they are observed (meta-analysis: Bradley, Lawrence, & Ferguson, 2018), and they consequently benefit from earning a reputation for generosity (see reviews in Barclay, 2010a, 2012a, 2015). Here we present a simple hypothesis: because protecting the environment is a form of cooperation, it will be maintained by many of the same factors that maintain other forms of cooperation — the same cost–benefit logic applies regardless of the particular domain of cooperation. In particular, because opportunities to earn a reputation are so important in supporting other forms of cooperation, we propose that they will also support environmentalism (see Sections 1.1 and 2).

There is currently some preliminary evidence that reputational benefits underlie (some) environmental behaviours. People are more pro-environmental when their actions are known to others, both in laboratory studies (e.g., Griskevicius et al., 2010; Milinski et al., 2006) and in field experiments (reviewed by Kraft-Todd, Yoeli, Bhanot, & Rand, 2015). For example, people buy recognizably hybrid cars like Toyota Priuses more often in locations with pro-environmental norms, but the same is not true for less conspicuous hybrids like Toyota Camry Hybrids (Sexton & Sexton, 2014). Even cues that imply reputational opportunities, such as photos of eyes which trigger a sense of being watched, can increase pro-environmental behaviours like cleaning up litter (Ernest-Jones, Nettle, & Bateson, 2011; Francey & Bergmüller, 2012) and donating money to environmental organizations (Ekström, 2011; Keller & Pfattheicher, 2011). However, if reputational benefits are an ultimate functional explanation for a pro-environmental psychology, then cooperators must benefit from being observed, and it is currently unknown whether people do actually benefit from being seen to protect the environment.

While we are arguing that reputational benefits underlie environmental behaviours, our argument does not require that all environmental behaviours are observed. A pro-environmental psychology will cause both public and private pro-environmental behaviours. As long as the net benefits for public environmentalism outweigh the costs of private environmentalism – on average – then reputational benefits will cause pro-environmental sentiment to proliferate.

1.1. Three types of reputation that could foster environmentalism

What are the potential reputational benefits of environmentalism? First, protecting the environment could be rewarded via indirect reciprocity (Milinski et al., 2006). Under indirect reciprocity, cooperative people get rewarded by observers: cooperators are more likely to receive help when they themselves need help (Nowak & Sigmund, 2005). Those who help the cooperators in turn get a good reputation and are rewarded by others. In this way, helpers have their help reciprocated indirectly, i.e., from people other than the recipients of the help. With environmentalism, those who protect the environment could be rewarded by those who observe it.

Second, incurring the costs of environmentalism could demonstrate one's wealth (Griskevicius, Cantu & van Vugt, 2012). Under this hypothesis, wealthy people are best able to afford the costs of environmentalism. Everyone gains reputational benefits from being seen to be "green", because audiences infer that the actor is wealthy. However, the benefits from being seen to be "green" are only worth gaining for those who really are wealthy enough to afford to donate to environmental causes or purchase sustainable products. For poorer people, the cost to them outweighs the benefits.

Our proposal represents a third, novel way that reputational benefits can support environmentalism: protecting the environment could be a signal of one's willingness to cooperate with others, such that observers trust and cooperate more with those who protect the environment (Barclay, 2012a). Those who are willing to cooperate with others will be more likely to protect the environment, because doing so helps everyone who shares that environment. Indeed, past work shows that people who value the environment are more likely to cooperate with others (Kaiser & Byrka, 2011; Sussman, Lavalle, & Gifford, 2016) and to have prosocial value orientations (Cameron, Brown, & Chapman, 1998; Gärling, Fujii, Gärling, & Jakobsson, 2003; Groot & Steg, 2008; Joireman, Lasane, Bennett, Richards, & Solaimani, 2001; van Vugt, Meertens, & van Lange, 1995; but see Joireman, van Lange, & van Vugt, 2004). As such, observers can then use a person's pro-environmental behaviour as a cue of that person's willingness to cooperate, and benefit from trusting that person accordingly. This trust provides an incentive for everyone to act more environmentally, and by doing so they are actively signaling their willingness to cooperate (for the distinction between cues and signals in this context, see the Supplementary Material).

Under this signaling hypothesis, everyone benefits from being seen to protect the environment, but it is only worth the cost of environmentalism for those who actually value others' welfare and intend to stay and cooperate for the long term. Thus, those who value others' welfare will be pro-environmental because the long-term benefits to them outweigh the costs. By contrast, for those who do not intend to cooperate, the costs of protecting the environment outweigh the short-term benefits of "suckering" someone into trusting. Thus, those who do not value others' welfare, or who do not intend to stay and cooperate for the long term, will not be pro-environmental because the costs outweigh the benefits to them of appearing environmental. This hypothesis is a direct application of theories that have been applied to other diverse forms of cooperation (André, 2010; Smith & Bliege Bird, 2005), including food sharing (Bliege Bird, Ready, & Power, 2018), romantic relationships (Bolle, 2001), charitable giving (Brekke, Hauge, Lind, & Nyborg, 2011; Hauge, Brekke, Nyborg, & Lind, 2019; Keser, 2003), religion (Sosis, 2004), costly apologies (Ohtsubo & Watanabe, 2009), and honesty in business (Frank, 2004; Pfeiffer, Tran, Krumme, & Rand, 2012). Here we apply those same principles to environmental sentiment.

Our signaling theory of environmentalism and cooperative intent – and the theory that environmentalism signals' one's wealth – are both part of "biological markets theory", wherein organisms advertise whatever valuable traits make them more desirable as partners (Barclay, 2013; Noë & Hammerstein, 1994, 1995). Desirable traits are those which indicate one's *ability* to benefit others or one's *willingness* to do so. The value of a trait depends on how useful that trait is in that social environment (e.g., physical coordination is desirable in teammates when sports are important), how diagnostic a given action is of some underlying quality (e.g., spinning a basketball is a better signal of coordination than is fruit-picking), and the supply and demand for that quality. To apply biological markets theory (Barclay, 2013, 2016): environmentalism advertises one's willingness to benefit others (i.e., one's cooperative intent).

Given that observers will preferentially associate with those who protect the environment, this creates an incentive (conscious or not) to actively broadcast one's environmentalism instead of keeping it private, and to act more environmentally when observed. Furthermore, given that organisms are in implicit competition to be chosen as social partners, people should actively compete to be more environmental than others, just as they compete to help in other ways (Barclay, 2004, 2013; Barclay & Willer, 2007; Raihani & Smith, 2015; Roberts, 1998; Sylwester & Roberts, 2010; van Vugt, Roberts & Hardy, 2007). Despite this competition making *everyone* act more environmentally, environmentalism will still be correlated with cooperative intent (or wealth) because people's relative rankings are

preserved. That is, the best cooperators are still the most environmental, and the worst cooperators are still least environmental; the overall levels of environmentalism are simply higher. For a discussion of how signals maintain their information value despite competition to signal more, see Biernaskie and colleagues (Biernaskie, Perry, & Grafen, 2018) and the Supplementary Material.

1.2. Predictions

Our broad proposition is that people use (consciously or not) environmentalism as a signal of cooperative intent within a biological market for social partners, and audiences (consciously or not) evaluate it as such. Based on this, we predict that the results observed for other signals of cooperative intent will also be observed in the domain of environmentalism (e.g., competitive altruism in Barclay & Willer, 2007). These predictions – although already tested in other forms of cooperation (e.g., Barclay & Willer, 2007; Feinberg et al., 2014; Hauge et al., 2019; Milinski et al., 2002; Wu et al., 2016) – would represent new findings about environmentalism and would thus support our hypothesis. Here we focus on the following seven predictions that our signaling hypothesis makes, noting that some but not all are also made by the other functional hypotheses about reputational benefits and environmentalism (indirect reciprocity and signaling wealth).

First, we predict that people who do more for the environment are more likely to cooperate with others (Prediction 1a) and have less selfish personalities (Prediction 1b). These correlations are necessary to demonstrate signaling: a signaling system is only stable if those who signal more (e.g., act more environmentally) are also higher in the trait being signaled (e.g., cooperative intent). Second, we predict that those correlations (i.e., 1a & 1b) will be mediated by other traits relevant to cooperation, such as one's valuation of others or of the future, such that the correlations are reduced when controlling for valuation of others or another relevant trait (Prediction 2).

Third, because our hypothesis is about signaling, we predict that people will be more environmental when their actions are observed by others instead of anonymous, i.e. when they have an opportunity to earn a reputation (Prediction 3). Fourth, we predict that partner choice will cause people to compete to protect the environment more than others do (Prediction 4). This predicted competition is unique to biological markets theory (and variants thereof, e.g., Roberts, 1998), and is not predicted by traditional explanations for cooperation based on direct or indirect reciprocity alone (i.e., without partner choice) (Axelrod, 1984; Nowak & Sigmund, 2005). Following the same logic, our fifth prediction is that choosers will prefer partners who give more to the environment, such that those who give tend to benefit by an increased likelihood of being chosen (Prediction 5). This kind of partner choice would be consistent with (and a logical extension of) indirect reciprocity, but is not explicitly required by that theory. Sixth, we predict that those who give will also benefit from increased cooperation within any new relationship (Prediction 6).

Finally, we make another prediction that is unique to signals of cooperative intent: the correlation between environmentalism and subsequent cooperation will be weakened (but not eliminated) when there are more incentives to appear environmental (Prediction 7). When people have more incentives to appear environmental, they are more likely to act environmentally, even if they have low intention to cooperate in the future. When less cooperative people start acting environmentally, it will weaken the correlation between environmentalism and cooperation, but will not eliminate that correlation (see (Biernaskie et al., 2018), for signal honesty despite an escalation of signaling). Hence Prediction 7.

We first show in a mathematical model how environmentalism could demonstrate one's likelihood of cooperation. We then present Studies 1 and 2, which were online studies to test Predictions 1a and 2, and are an important proof of concept of our model. Finally, we present Studies 3 and 4, which used an experimental economic game with subsequent exploratory questionnaires. These are the most important studies, because in addition to testing Prediction 1, they also test Predictions 3–7.

The other reputation-based theories of environmentalism could also make some of these predictions, given that all these theories are about reputational benefits, but each makes a unique combination of predictions, and no other theory except for signaling cooperation makes all of the predictions (see Table 1). Indirect reciprocity predicts that environmentalism will correlate with cooperation (Prediction 1a) and personality type (Prediction 1b), and could be extended to predict that concern for others mediates those correlations (Prediction 2). Indirect reciprocity also predicts that people will be more environmental when observed (Prediction 3) and that people who act more environmentally will receive more cooperation from others (Prediction 6). Furthermore, while not explicitly making the prediction, indirect reciprocity theory could implicitly say that people who act environmentally are chosen more often as partners (Prediction 5). However, it explicitly does not predict that people will act even more environmentally when competing to be chosen (i.e., silent about Prediction 4), nor that the environmentalism-cooperation link is weaker when there are more incentives to appear environmental (i.e., silent about Prediction 7).

By contrast, if environmentalism functions as a signal of wealth (Griskevicius, Cantu & van Vugt, 2012), then we should make several different predictions. Like our cooperativeness-signaling hypothesis, the wealth-signaling hypothesis is ultimately derived from biological markets theory (Barclay, 2013; Noë & Hammerstein, 1994, 1995), so it predicts that people will be more environmental when observed (Prediction 3), will compete to be more environmental than others (Prediction 4), and will be preferentially chosen as partners (Prediction 5). However, the wealth-signaling hypothesis makes no predictions about cooperation. Therefore, it does not necessarily predict that environmental people will receive more cooperation from partners (i.e., silent about Prediction 6). Furthermore, it explicitly does not predict a correlation between environmentalism and cooperation or personality type (i.e., silent about Predictions 1a & 1b), nor any mediation or weakening of that correlation (i.e., silent about Predictions 2 and 7, respectively). Instead, the wealth-signaling hypothesis explicitly predicts a correlation between environmentalism and wealth or salary, and possibly other traits that could be a proxy for wealth (e.g., intelligence). If there is a correlation between wealth and environmentalism, then audiences can infer one's wealth from one's environmentalism, which then gives people an incentive to display their environmentalism in order to signal that they are wealthy. By contrast, if there is no correlation between wealth and environmentalism, then audiences should not infer someone's wealth from their environmentalism, and environmentalism will not be used to signal wealth.

1.3. Testing the function of environmentalism, not the proximate psychological mechanisms

Before we introduce our studies, we must stress which level of analysis we are examining. Many psychological studies test the proximate psychological mechanisms underlying behaviour. For example, which specific norms, values, emotions, attitudes, and concerns motivate environmental behaviour (e.g., de Groot & Steg, 2008; Steg & Vlek, 2009; Steg et al., 2014; van der Linden, 2018)? Other psychological studies test the developmental mechanisms underlying behaviour. For example, what types of learning cause people to become pro-environmental (e.g., imitation vs. norm-following vs. reinforcement learning), and more broadly, how do nature and nurture interact to cause people to become pro-environmental? These are important questions, and are

well-studied for good reason. However, they are not the only kinds of questions one can ask.

Others ask instead about the ultimate functions underlying behaviour. Why do these proximate psychological mechanisms and developmental responses exist at all? Why has not everyone learned to love (or have an innate love for) destroying the environment instead of protecting it? Why does a pro-environmental psychology arise in at least some people – what benefits does it bring? What payoffs have resulted in pro-environmental sentiment persisting – or even spreading – instead of being eliminated due to its costs? In other words, what is its function? This kind of question – ultimate function – is much less-studied for environmentalism than questions at the other levels of analysis.

Here, we are investigating the function of environmentalism, which is complementary to questions of proximate psychological mechanisms and development (Barclay, 2012b; Scott-Phillips et al., 2011; Tinbergen, 1968). We propose that (some) people possess a pro-environmental psychology because acting on one's pro-environmental sentiments can increase others' opinions of oneself (i.e., one's reputation). In other words, acting pro-environmentally brings positive social consequences when others know about one's environmentalism. This does not mean that environmentalists are consciously concerned about others' opinions — they may be, but need not be. We are agnostic about which proximate mechanisms are involved. Instead, we simply mean that when the reputational benefits are greater, pro-environmental psychologies will be more prevalent or more active, and that people who behave more environmentally will benefit from doing so (regardless of whether they intend to benefit). It is these kinds of questions that we test.

2. Mathematical models: Why do pro-environmental people cooperate more with partners?

We argue that environmentalism is associated with people's willingness to cooperate with others, such that audiences can infer a person's cooperativeness from that person's environmentalism. But *why* would environmentalism be correlated with cooperation? What traits would cause the same person to both protect the environment and cooperate directly with others? Here we present a mathematical model to investigate this. Mathematical models – even simple ones – are useful as formal proofs of concepts, formalizations of the assumptions, and as means of making precise quantitative predictions (for the importance of mathematical models, see Muthukrishna and Henrich (2019), Servedio, Brandvain, Dhole, Fitzpatrick, Goldberg, Stern, Van Cleve, and Yeh (2014)).

Protecting the environment is a public good that benefits multiple parties, so protecting the environment could be a cue of one's concern for all those affected. (See Supplementary Material for discussion of whether environmentalism may be a cue rather than a signal.) This argument is similar to other models of gift-giving as a signal of cooperative intent (e.g., Andreoni & Bernheim, 2009; Barclay, 2013; Fehrler & Przepiorka, 2013), but applied specifically to environmentalism. Affected audiences can infer some minimum level of concern, i.e., the environmentalist values the affected audience at least enough to provide them the environmental public good, and possibly values them more than that. This "minimum concern" lets members of the affected audience predict whether the environmentalist will cooperate with them. In Supplementary Material, we present a mathematical model showing that affected audiences can infer that an environmentalist will cooperate with them when, at the least:

$$b_P(c_e - b_e)/nb_e > c_P \qquad \text{(Inequality 1)}$$

where c_e is the cost of providing an environmental good which benefits n recipients by b_e each, and c_P and b_P are the costs and benefits of giving and receiving help to audience members in a Prisoner's Dilemma, respectively. Thus, expensive environmental goods (high c_e) that provide low benefit (low b_e) to few people (low n) indicate the most concern for those people, and thus demonstrate the environmentalist's high willingness to cooperate with the beneficiaries.

Instead of being a cue of one's valuation of others, environmentalism could be a cue of one's time horizons, given that reciprocity requires a long "shadow of the future" (André, 2010; Axelrod, 1984; Curry, Price, & Price, 2008; Jordan, Hoffman, Bloom, & Rand, 2016). We created a second mathematical model to test this alternate hypothesis, but we present it only in Supplementary Material because Studies 1 and 2 did not provide empirical support for it.

3. Studies 1 & 2

Our mathematical models show that observers can use someone's environmentalism as a cue of that person's concern for others (Ineq. (1)) or their time horizons (Supplementary Material). We conducted two online studies (Studies 1 & 2) on Amazon Mechanical Turk (MTurk) to test these two mathematical models: environmentalism as a cue of one's concern for others versus a cue of one's time horizons. Thus, our model predicts that environmentalism will be correlated with cooperation (Prediction 1a). If this correlation exists because people's concern for others underlies both environmentalism and cooperation (Ineq. (1)), then concern will mediate the correlation between environmentalism and cooperation (Prediction 2). A similar argument holds for time horizons.

To test these predictions, we asked workers on crowdsourcing website Amazon Mechanical Turk ("MTurk") to play an online one-round Continuous Prisoner's Dilemma (CPD) where they were paired with another anonymous MTurk worker, and each could give up to US$0.50 to their partner, with any money given being doubled by the experimenter before being received. Participants then filled out questionnaires assessing their environmental behaviours, concern for others, and salary. Study 1 used an environmental title in recruitment; Study 2 was a replication using a neutral title. Studies 1–2 are a "proof of concept" to test our first two predictions.

3.1. Methods for Studies 1 & 2

3.1.1. Participants.

Participants were 226 (Study 1) and 309 (Study 2) people recruited from Amazon Mechanical Turk (MTurk) for a short study. MTurk is an online crowdsourcing site where workers ("Turkers") complete small tasks for small amounts of money, and is increasingly used in behavioural research (for details, see: Rand, 2012). Of those participants, 218 (Study 1) and 301 (Study 2) completed all questionnaires we needed for a mediation analysis and correctly responded to our question to weed out non-human bot responders ("what colour is the sky on a clear day with no clouds?"). Participants in Study 1 reported a median salary of US$30,000 (interquartile range: US$15,000-$50,000), and participants in Study 2 reported a median salary of US$32,000 (interquartile range: US$18,500-US$54,500); about 6% of participants in each study reported a salary over US$100,000.

3.1.2. Incentives and recruitment framing.

Studies 1 and 2 differed only in the baseline pay and the title that was displayed when recruiting MTurk workers: "Greener Than Thou" (Study 1) or "Decisions and Questionnaires" (Study 2). Participants received US$0.30 (Study 1) or US$0.50 (Study 2) base pay for participating (MTurk rates changed between the two studies), plus bonuses based on the Prisoner's Dilemma (see below). These rates were typical for MTurk for a short experiment (less than 10 min) at the time of the experiments.

3.1.3. Measures

3.1.3.1. Prisoner's Dilemma. Participants played an online one-round Continuous Prisoner's Dilemma (CPD). In the CPD, each participant was paired with another MTurk worker. Each member of the pair received US$0.50 and could give any amount to their partner (i.e. it is "continuous" because this giving is a non-binary decision). The experimenter doubled any donations, such that the partner received twice the dollar amount given, e.g., it cost a participant $0.10 to confer $0.20 on their partner. This game is a Prisoner's Dilemma because it is costly to give up money (the money-maximizing strategy is to give nothing and hope that the other person gives), but mutual cooperation pays better than mutual non-cooperation ($1.00 vs. $0.50, respectively). These decisions were anonymous: no one knew participants' identities, and the experimenter only knew participants' MTurk ID, which is a long string of letters and digits. After the CPD, participants filled out the following questionnaires, and received their base pay and CPD pay shortly after participating.

3.1.3.2. Social Value Orientation (SVO). To measure concern for others, we used Social Value Orientations (SVO), which are a reliable predictor of how people value others relative to themselves (Balliet, Parks, & Joireman, 2009). SVOs are stable personality traits with good test-retest reliability (e.g., Murphy & Ackermann, 2014). To assess SVO, we used the "slider" measure of SVO (Murphy, Ackermann, & Handgraaf, 2011), which produces a continuum of SVO instead of discrete categories. Participants made a series of six decisions about how to divide hypothetical money between themselves and a hypothetical partner (e.g., 85 for Self & 85 for Other vs. 87 for Self and 81 for Other). For each question, there are nine options on a scale that varies in terms of personal gain, collective gain, and/or relative advantage over one's partner. Participants use a slider bar to indicate which of the nine options they prefer. The experimenter uses participants' responses to calculate an "SVO angle" for each participant, which measures how much each person values their own earnings against the earnings of others. Whereas the traditional SVO scale produces a categorical classification of each participant (e.g., proself vs. prosocial Balliet et al., 2009), the SVO slider measure produces a continuous measure: "altruists" have very high SVO angles, "prosocials" have moderately high SVO angles, "individualists" have low SVO angles, and "competitors" have very low SVO angles.

3.1.3.3. Time horizons. We measured participants' time horizons with 20 questions on their preferences for a small of money soon or more money later (e.g., $41 tomorrow vs. $51 in 33 days). We counted the number of times they chose the delayed but larger option (Griskevicius et al., 2012).

3.1.3.4. Environmental behaviour scale. We assessed pro-environmental behaviour with Casey and Scott's (2006) environmental behaviour scale. Participants responded on a 1 ("never") to 4 ("always") Likert scale to questions like "where possible, I buy products made from recycled materials as opposed to those not made from recycled materials" or "when cleaning my teeth I turn off the tap rather than leaving it run". Such behaviours may be cheap to perform once, but performing them repeatedly over time carries higher costs and may be good for building trust with long-term partners (Bliege Bird et al., 2018). A meta-analysis by Kormos and Gifford (2014) shows that self-reported pro-environmental behaviours correlate with objectively measured pro-environmental behaviours at $r = .46$.

3.1.3.5. Other measures. To test the hypothesis that environmental behaviour signals one's wealth (e.g. Griskevicius, Cantu & van Vugt, 2012), we asked participants to give their estimated annual income in US dollars. For exploratory purposes, we also asked participants in Study 2 to complete Snyder and Gangestad's (1986) Self-Monitoring scale, which includes yes/no questions like "In different situations and with different people, I often act like very different persons". Studies 1 and 2 were approved by the Research Ethics Board at the University of Guelph.

3.2. Results of Studies 1 & 2

3.2.1. Was environmentalism correlated with cooperation (Prediction 1a) and valuation of others (Prediction 2)?

Yes, in both studies independently. As predicted, cooperation in the Prisoner's Dilemma was positively correlated with scores on the environmental questionnaires in both Study 1 and Study 2 ($r_{224} = .26$ and $r_{301} = .11$; 95% CIs [.13, .38] and [.00, .22]; $p < .001$ and $p = .054$, respectively, combined $r_{527} = .17$, 95% CI [.09, .25], $p < .001$). The correlation was somewhat stronger in Study 1 than Study 2 ($z = 1.70$, 2-tailed $p = .09$). Cooperation was also correlated with Social Value Orientation in both studies ($r_{219} = .35$ and $r_{299} = .33$, 95% CIs [.24, .47] and [.24, .44], both $ps < .001$), and environmentalism was also correlated with Social Value Orientation in both studies ($r_{217} = .24$ and $r_{304} = .21$, both $ps < .001$).

As predicted, after controlling for Social Value Orientation, environmentalism was less correlated with cooperation in both Study 1 and Study 2 (partial correlation $r_{215} = .18$ and $r_{298} = .04$, $p = .008$ and .48, respectively). To test whether the causation goes the other way, we note that controlling for environmentalism did not reduce the correlation between Social Value Orientation and cooperation, in either Study 1 or Study 2 (partial correlation $r_{215} = .32$ and $r_{298} = .32$, respectively, $ps < .001$). Sobel's test of mediation showed that Social Value Orientation mediated the environmentalism-cooperation link in both Study 1 (test statistic 2.70 ± s.e. 0.00057, $p = .007$) and in Study 2 (test statistic 3.00 ± s.e. 0.00051, $p = .003$), with 30.8% and 62.3% mediation, respectively. This supports the hypothesis we developed in the mathematical model that environmentalism is a cue of one's cooperative intent in part because it reflects one's concern for others. However, the incomplete mediation suggests that other factors may be involved.

To show the robustness of our mediation analyses, we ran bootstrapping mediation models with 5000 bootstrap resamples using the INDIRECT macro in SPSS 26.0 (Preacher & Hayes, 2008). There was a total effect of environmentalism on cooperation in both Study 1 and Study 2 (total effects: B = 0.48 ± s.e. 0.13, $t_{214} = 3.77$, $p = .0002$; and B = 0.24 ± s.e. 0.12, $t_{301} = 1.90$, $p = .059$, respectively). In both studies, the direct effect of environmentalism on cooperation was lower than the total effect (direct effects: B = 0.33 ± s.e. 0.12, $t_{215} = 2.66$, $p = .0083$; and B = 0.09 ± s.e. 0.12, $t_{301} = 0.71$, $p = .48$, respectively), because Social Value Orientation mediated the relationship between environmentalism and cooperation in both Study 1 and Study 2 (indirect effect B = 0.15 ± s.e. 0.05, 95% CI [0.07, 0.27]; and B = 0.15 ± s.e. 0.05, 95% CI [0.05, 0.26]; respectively). These bootstrapped analyses support our original analysis: environmentalism seems to be a cue of cooperative intent because it reflects concern for others.

To assess the direction of mediation, Pieters (2017) recommends basing it on theory and prior research rather than reverse causation, and there are multiple theoretical reasons to support our interpretation. First, our mathematical model shows that valuation of others (e.g., SVO) causes people to cooperate more. Second, SVO is a stable personality trait with good test-retest reliability (Balliet et al., 2009; Murphy et al., 2011) whereas cooperation is more sensitive to context, so it is less likely that the latter causes the former. Third, SVO is a measure of people's preferences, and people's actions are generally assumed to be influenced by their preferences for different outcomes (e.g., prospect theory: Tversky & Kahneman, 1992), more so than vice versa. Nevertheless, given that some mediation does occur in the other direction in Study 1 (but not Study 2, Table S2), it is likely that any of these cooperative acts or traits is useful for predicting the others.

3.2.2. Were other correlations or mediations important?

No, in neither study. Environmentalism was somewhat positively correlated with time horizons in Study 1 and 2 (r_{208} = .19 and r_{299} = .09, 95% CIs [.06, .32] and [−.02, .20], p = .008 and .11, respectively), but time horizons were not strongly correlated with cooperation (r_{206} = .08 and r_{299} = .04, 95% CIs [−.06, .21] and [−.07, .15], p = .23 and .48, respectively). As such, time horizons cannot mediate the relationship between environmentalism and cooperation.

Salary was not correlated with either environmentalism (Study 1: r_{215} = −.09, 95% CI [−.22, .04], p = .19; Study 2: r_{298} = −.00, 95% CI [−.11, .11], p = .97) or with cooperation (Study 1: r_{215} = .03, [−.10, .16], p = .67; Study 2: r_{298} = −.01, 95% CI [−.12, .10], p = .94). Self-Monitoring did not correlate with anything (all r < .08, all p > .19). Table S1 includes all correlations in both studies.

3.3. Discussion of Studies 1 & 2

As predicted, Studies 1 and 2 found that participants' environmentalism predicted their likelihood of cooperating with others (Prediction 1a). Thus, Studies 1 and 2 provide a preliminary proof of concept, using everyday environmental behaviours (for the validity of self-reported environmental behaviours, see the meta-analysis by Kormos and Gifford (2014)). Such behaviours have low cost if performed once, but higher costs if performed regularly.

In addition, Studies 1 and 2 show that concern for others mediates the environmentalism-cooperation link (Prediction 2), although the mediation is incomplete. This supports the idea that environmentalism is a cue of cooperation because it is informative about people's concern for others, as predicted by our mathematical model (Model 1). By contrast, Model 2 was not supported: time horizons did not mediate the environmentalism-cooperation link. Furthermore, salary was uncorrelated with either environmentalism or cooperation; range restriction was not a problem because the participants in both studies had a wide range of salaries. If salary does not correlate with a given environmental behaviour, then people are unlikely to use that kind of environmentalism to signal their wealth, because audiences will not infer wealth when they observe that environmentalism. (Note: this holds for low-cost environmentalism, such as recycling, whereas a correlation with salary is more likely for expensive environmental acts, e.g., buying a Prius). Finally, Study 2 replicated Study 1 without an environmental framing (albeit not as strongly), suggesting that the effect exists even among people who are not primed to think about environmentalism.

4. Studies 3 & 4

Studies 1 and 2 showed that online participants who do more for the environment were more cooperative in experimental games, in part mediated by concern for others (Predictions 1 & 2). Studies 3 and 4 were in-lab studies that tested our remaining predictions: whether people donate more money to the environment when observed (Prediction 3) and especially when competing over partnerships (Prediction 4); whether high donors are chosen more often as partners (Prediction 5) and receive more money from others (Prediction 6); whether environmental donations are correlated with one's cooperation towards a partner (Prediction 1a and 1b); and whether that correlation is reduced with the incentives to appear environmental (Prediction 7). Such results have been found for other types of non-environmental cooperation (e.g., Barclay & Willer, 2007); here we test whether they apply to environmentalism as well.

We randomly assigned participants into groups of three: Players A and B could donate money to an environmental charity (the Sierra Club), and then one of them could play a cooperative game (Continuous Prisoner's Dilemma) with the third player (Player C). There were three experimental conditions. Player C either: (a) did not observe A and B's donations and was randomly assigned to partner with either A or B (Random/Anonymous condition); (b) observed A and B's donations and was randomly assigned to partner with A or B (Random/Knowledge condition); or (c) observed A and B's donations and could choose either A or B for the Continuous Prisoner's Dilemma (Choice/Knowledge condition). Given the importance of replication in science, Study 4 replicated Study 3 using an environmentally framed recruitment poster instead of a neutrally framed poster.

4.1. Methods for Studies 3 & 4

4.1.1. Participants, incentives, anonymity, and recruitment framing.

We recruited participants of various ethnic backgrounds from the University of Arizona community via posters and university listservs. Participants received US$5 for participating plus US$1 for every 5 lab dollars (henceforth L$) earned in the experiment. Participants received code numbers, and we used a two-experimenter procedure to maintain anonymity of participants' decisions: one experimenter knew participants' identities and code numbers but not their decisions, and the other experimenter knew the decisions and code numbers but not who was who (see Supplementary Material). Participants could associate others' decisions with a code number, but not the person's actual identity.

Study 4 was identical to Study 3 except that participants were recruited with a specifically environmental framing. The recruitment posters for Study 3 had invited participants to "receive money for participating in an experiment on decision-making" because "we are seeking participants for studies of people's decision-making". By contrast, the recruitment posters for Study 4 specifically mentioned environmentalism by inviting participants to "receive money for participating in an experiment on environmental decisions" because "we are seeking participants for studies of people's decisions regarding environmental conservation" (see Supplementary Material). The framing for recruitment – regular or pro-environmental – was the only difference between Studies 3 and 4. Participants in Study 4 were similar to those in Study 3, except for being significantly less religious, marginally more pro-environmental, and valuing the Sierra Club significantly more (see Supplementary Material Table S7).

A power analysis (using the program G*Power) showed that we needed 36 participants in each study to have a 90% chance of detecting a medium-sized effect (f = .25) with our design. In Study 3, 25 males and 29 females (mean age 25.3 years ± s.d. 10.6 years) participated and earned an average of US$13.00 ± s.d. US$4.02, while in Study 4, 21 males and 33 females (mean age 25.6 years ± s.d. 11.8 years) participated and earned an average of US$13.11±s.d. US$2.65. These methods were approved by the University of Arizona Institutional Review Board.

4.1.2. Procedure.

We used a modified version of Barclay and Willer's (2007) within-subjects choice task. Nine people took part in each session, assigned to three groups of three people. In Part 1, two members of each group ("A" and "B") were each given L$10 and could each donate any amount (in multiples of L$0.01) to the Sierra Club, a well-known environmental charity in the United States (information on the Sierra Club was available on everyone's desks). These donations are the main dependent variable, and the Sierra Club received twice the amount that was donated (after converting from L$ to US$). The third person in each group ("C") sat out.

In Part 2, that third member (C) was paired with either A or B and played a Continuous Prisoner's Dilemma (henceforth CPD) with her (Roberts & Renwick, 2003). Each member of the pair playing the CPD received L$10 and could give any amount to her partner. The experimenter doubled any amounts given, such that both partners were better off if both gave, but giving was personally costly. We had three experimental conditions for the CPD. Player C was either: (a) randomly paired with either A or B and would not know how much

each had donated to the Sierra Club in Part 1 ("Random/Anonymous" condition); (b) randomly paired with either A or B and was told how much each had donated to the Sierra Club ("Random/Knowledge" condition); or (c) was told how much each had donated to the Sierra Club and could choose whom to pair with for the CPD ("Choice/Knowledge"). Everyone was told in advance about the game structure (see Supplementary Material for instructions), and knew when they were making decisions for each condition. As such, contrasting the Random/Anonymous and Random/Knowledge conditions tests whether people would donate more to the environment to gain a good reputation (which could then benefit them in the CPD). Contrasting the Random/Knowledge and Choice/Knowledge conditions tests whether people will compete to donate *more* than someone else in order to be chosen for the CPD. The latter is the crucial comparison for competitive helping, and is uniquely predicted by biological markets theory and variants thereof.

We used a within-subjects design to factor out individual differences in cooperation: each participant experienced (and was paid for) each of the three experimental conditions once, each within a completely different group of three players in each condition. Order of conditions was counterbalanced across sessions. Participants made all Part 1 decisions before making any Part 2 decisions, and did not receive feedback on any Part 1 until right before Part 2, so the results of one condition could not affect the Part 1 decisions of other conditions. We elicited A's and B's Part 2 decisions using the "strategy method" (Fehr & Fischbacher, 2004) to maximize the amount of data we could collect: participants indicated how much they would give to their partner if they were involved in the CPD, and we implemented the decisions of the person who was selected. Feedback for Part 2 was only provided after all decisions had been made. There was an error in recording the Part 2 decisions in one session, so we scheduled another session with new participants to replace it; the results are qualitatively unchanged if the original session is used instead of the replacement. Participants completed exploratory questionnaires after the experiment (see Supplementary Material).

4.1.3. Statistical analysis.

We calculated 95% confidence intervals for within-subjects data using the recommendations of Jarmasz and Hollands (2009) and Loftus and Masson (1994). To analyse Studies 3 and 4 separately (Supplementary), for each study we used a Repeated-Measures GLM in SPSS 23 with the three experimental conditions as within-subject factors. We corrected for violations of sphericity using a Greenhouse-Geisser correction. Results are qualitatively similar – and often stronger – if non-parametric tests are used (see Supplementary Material, Table S5, S6). Unless otherwise noted, all contrasts were planned a priori based on Barclay and Willer (2007) and use two-tailed tests and effect sizes generated by SPSS. As in Barclay and Willer (2007), sample sizes were pre-determined by the need to have six 9-person sessions in order to counterbalance the order of conditions; Studies 3 and 4 thus had 54 participants each (36 as Player A or B, 18 as Player C), which our power analysis shows gave us 90% power to detect a medium-sized effect ($f = .25$).

Given that Studies 3 and 4 produced qualitatively similar results that were significant independently of each other, we combined Studies 3 and 4 to increase our sample size and tighten our confidence intervals. Thus, here in the main text we present a 2 (Framing) × 3 (Type of Reputation) mixed between- and within-subjects GLM of amounts donated to the Sierra Club. In the Supplementary Material we present an analysis of Study 3 separately from Study 4, so interested readers can see how Study 4 directly replicates Study 3.

4.2. Results of Studies 3 & 4

As predicted, our main analysis shows a main effect of experimental condition on donations ($F_{2,140} = 27.70$, $p < .001$, $\eta^2 = .284$). In the exploratory part of the analysis, framing had no main effect ($F_{1,70} = 1.19$, $p = .28$, $\eta^2 = .017$), nor did it interact with experimental condition ($F_{2,140} = 1.80$, $p = .18$, $\eta^2 = .025$). Sections 4.2.1 and 4.2.2 conduct planned analyses to test which experimental conditions were different.

4.2.1. Did people donate more when their donations were known to others (Prediction 3)?

Yes, and this was independently significant in both Study 3 and Study 4 (Fig. 1). Participants donated more of their endowment in the Random/Knowledge condition than in the Random/Anonymous condition (combined analysis: 45.5% [95% CI 37.7–53.2] vs. 32.2% [95% CI 24.2–40.1], respectively, $F_{1,70} = 28.09$, $p < .001$, $\eta^2 = 0.286$, mean difference 13.3% [95% CI 8.3–18.4]). The different framings of Study 3 and 4 produce qualitatively similar results: there was no main effect of study framing (Study 3: 36.0% [95% CI 25.5–46.6] vs. Study 4: 41.6% [95% CI 31.0–52.2], $F_{1,70} = 0.56$, $p = .46$, $\eta^2 = 0.008$), nor did framing interact with the observability of donations ($F_{1,70} = 2.26$, $p = .14$, $\eta^2 = 0.031$).

4.2.2. Did people compete to donate more to the environment (Prediction 4)?

Yes, and this was independently significant in both Study 3 and Study 4 (Fig. 1). Participants donated more in the Choice/Knowledge condition than in the Random/Knowledge condition (combined analysis: 52.5% [95% CI 45.0–60.0] vs. 45.5% [95% CI 37.7–53.2], respectively, $F_{1,70} = 9.40$, $p = .003$, $\eta^2 = 0.18$, mean difference 7.0% [95% CI 2.5–11.6]). The different framings of Study 3 and 4 produce qualitatively similar results: there was no main effect of study framing (Study 3: 43.7% [95% CI 33.4–53.9] vs. Study 4: 54.3% [95% CI 44.0–64.6], $F_{1,70} = 2.13$, $p = .15$, $\eta^2 = 0.030$), nor did framing interact with the observability of donations ($F_{1,70} = 0.32$, $p = .58$, $\eta^2 = 0.004$).

4.2.3. Did observers choose the person who gave more to the environment (Prediction 5)?

Yes, and this was independently significant in both Studies 3 and 4. Across both studies, Players A and B donated the same amount 7/36 times, and Player C chose the more generous person in 28 of the other 29 times (binomial $p < 0.0001$ both including and excluding ties; note that $N = 36$ because Player C could only choose partners in the Choice/Knowledge condition).

4.2.4. Did observers cooperate more with people who gave more to the environment (Prediction 6)?

Yes in the combined analysis, as long as the donations were known (i.e., Random/Knowledge and Choice/Knowledge conditions). As predicted, Player C tended to give more to Players A and B who had donated more to the Sierra Club: Random/Knowledge condition $r_{34} = .35$, $p = .04$, 95% C.I. [.02, .61]; Choice Knowledge condition $r_{34} = .39$, $p = .02$, 95% C.I. [.07, .64].

In the Random/Anonymous condition, we neither predicted nor found a positive correlation between how much A or B donated and how much she received from C ($r_{34} = -.36$, $p = .03$, 95% C.I. [−.62, −.04]). Any deviations from zero in the Random/ Anonymous condition must be Type I errors because C did not know what A and B donated. We present them only for posterity.

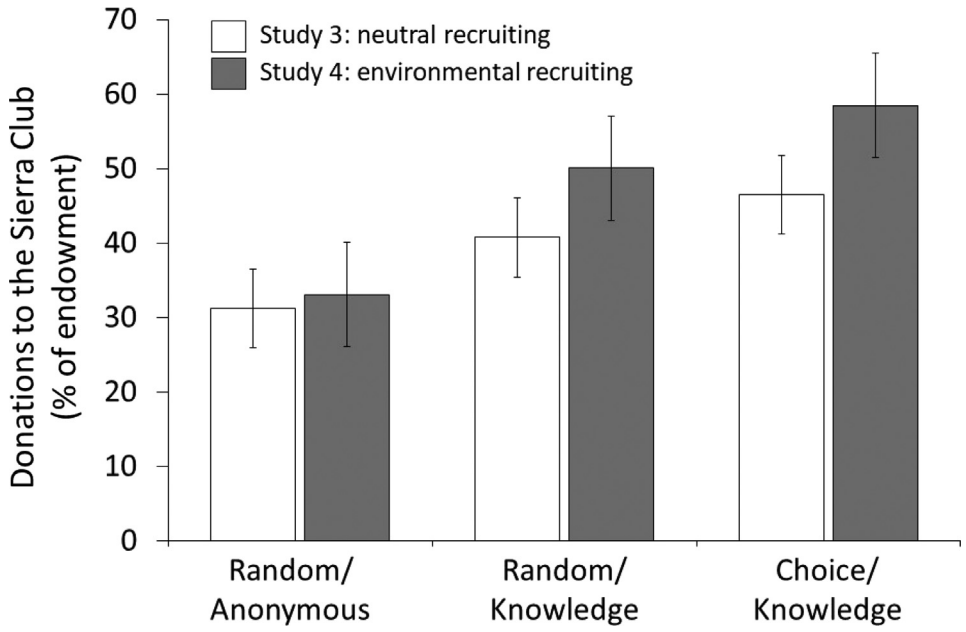

Fig. 1. Mean donations (±95% C.I.) to the Sierra Club in Study 3 (neutral recruiting, white bars) and Study 4 (environmental recruiting, shaded bars). Observers either did not know donors' decisions and were randomly assigned partners (Random/Anonymous), knew donors' decisions and were randomly assigned partners (Random/Knowledge), or knew donors' decisions and could choose partners (Choice/Knowledge). Within-subjects confidence intervals were calculated as recommended by Jarmasz and Hollands (2009) and Loftus and Masson (1994).

4.2.5. Were donations to the environment an honest signal of future cooperation (Predictions 1a and 7)?

Yes, and this was independently significant in both Studies 3 and 4: participants who donated more to the Sierra Club tended to give more to their partners in the Continuous Prisoner's Dilemma (Prediction 1a), regardless of the experimental condition or type of recruiting. The correlations in the combined analysis are: Random/Anonymous condition $r_{70} = .59$, $p < .001$, 95% CI [.41, .72]; Random/Knowledge condition $r_{70} = .65$, $p < .001$, 95% CI [.50, .77]; Choice/Knowledge condition $r_{70} = .49$, $p < .001$, 95% CI [.29, .65]. See Supplementary Material for separate analyses of Study 3 and 4, where the individual coefficients of determination (r^2) ranged from 0.15 to 0.61. As predicted, correlations were somewhat weaker in the Choice/Knowledge condition than in the Random/Knowledge condition (Prediction 7, see Supplementary Material).

An exploratory post hoc analysis revealed that these correlations were sometimes stronger when participants were recruited with regular framing (Study 3) than when they were recruited with environmentally-framed posters (Study 4). The differences in correlation are as follows: Random/Anonymous condition $z = 2.57$, $p = .01$; Random/Knowledge condition $z = 1.91$, $p = .056$; Choice/Knowledge condition $z = 1.13$, $p = .26$. This result is possibly because of selection biases or because the framing changed people's behaviour away from their normal inclinations. However, because it was a post hoc exploratory analysis and not always significant, readers should treat these differences in correlation with caution.

4.2.6. Exploratory questionnaires (Prediction 1b).

The post-experiment questionnaires from Studies 3 & 4 give us demographic and personality data which are of secondary interest compared to the main experiment, but also allow us to do post hoc tests of some additional hypotheses (see Supplementary Material). Most importantly, and supporting Prediction 1b: participants' scores on a measure of selfishness (Machiavellianism) were negatively correlated with their giving to the Sierra Club ($r_{68} = -.31$, 95% CI [−.51, −.08], $p = .009$) and to their partners ($r_{68} = -.24$, 95% CI [−.45, −.01], $p = .044$). Furthermore, a composite measure of liking the Sierra Club was positively correlated with giving to both the Sierra Club ($r_{69} = .34$, [.12, .53], $p = .003$) and to partners ($r_{69} = .27$, 95% CI [.04, .47], $p = .021$).

Participants' self-reported wealth, grade-point average, physical attractiveness, and physical strength were all uncorrelated with their giving. People scoring higher on self-monitoring tended to alter their donations more in response to opportunities for reputation (r_{67}'s = .22 to .33, depending on the comparisons, all p's < .07, see Supplementary Material), but no other variable consistently predicted responsiveness to reputational opportunities. There was no clear evidence that Player Cs' environmentalism predicted how much they rewarded donations to the Sierra Club (see Supplementary Material).

4.3. Discussion of Studies 3 & 4

The results from Studies 3 and 4 support all of the predictions they tested. Participants donated more to the environment when observed (Prediction 3), competed to donate more to the environment than others (Prediction 4), gave more money to people who donated to the environment (Prediction 5), and preferentially chose partners who donated more to the environment (Prediction 6). Furthermore, participants' donations were highly informative about donors' likelihood of cooperating with partners (Prediction 1a), though less so when participants had the biggest incentive to appear cooperative (Prediction 7). We note that Study 4 replicates all our predicted findings from Study 3, and with similar magnitude; see Supplementary Material for analyses of each study separately. Furthermore, Studies 3–4 find no correlation between wealth and either environmental donations or attitudes (Supplementary); this replicates Studies 1–2 with different measures of wealth and environmentalism.

We did not compare environmental behaviour with other prosocial behaviours, or indeed with any other positive behaviours without a social component. Thus, we cannot address whether environmental behaviours are more desired than other behaviours, or are used more in competition for partners. However, this is not our goal. According to biological markets theory (e.g., Barclay, 2013, 2016), people will

signal *any* trait that is valued by others, and audiences will value any trait that predicts people's willingness and ability to provide benefits. Environmentalism is just one of many ways in which people can signal their cooperative intent to compete over partners; our contribution is to show that it does indeed function as such a signal. The value of environmentalism (relative to other traits) will depend on what traits people need in their partners. Future studies should test how different traits (e.g., environmentalism vs. physical coordination vs. intelligence) are valued in different situations (e.g., cooperation game vs. sports game vs. quiz game), and should test the relative diagnostic value of each trait (e.g., environmentalism vs. moral reasoning as cues of cooperative intent).

5. General discussion

We present a mathematical model and four studies showing that people who do more for the environment are more likely to cooperate with others. This result generalized across environmental behaviours like recycling and conserving water (questionnaires in Studies 1–2), monetary donations to the environment (Studies 3-4), attitudes towards conservationist groups (Studies 3-4), and abstract environmentalism in a mathematical model that can represent any kind of pro-environmental behaviours. Furthermore, Studies 3 and 4 show that participants responded to environmentalism as if it were a signal of cooperation, by not only choosing more environmental people as partners but also cooperating more with them. Such responses make sense: choosers benefited from attending to environmental donations, given that environmental donations accounted for 15%–61% of the variance in amounts that people later gave to partners. While not perfect predictors of cooperative intent, these correlations make environmentalism informative as a basis for trust and partner choice. Our mathematical model and mediation analyses suggest that environmental behaviour demonstrates a concern for others, which may be why it predicts future cooperation.

Participants also responded to reputational opportunities: they donated more to an environmental charity (the Sierra Club) when their donations were known to potential social partners than when unknown, but they donated the most when competing to be chosen. This represents an escalation of environmental behaviour in response to competition over social partners, above and beyond that of simply "looking good". In other words, people competed to do more for the environment (see also van Horen, van der Wal, & Grinstein, 2018). This shows that reputation-based partner choice can amplify pro-environmental behaviour even more than can observation alone. Furthermore, those who acted more environmentally benefited from doing so. This shows that the principles of reputation (e.g., signaling, audience responses) that apply to other forms of cooperation also apply to environmentalism.

Our study is unique in the environmentalism literature in that it investigates both signalers and observers: people who act environmentally, and others who see that environmentalism and react to it. A signaling system is only stable if both parties benefit from the system (e.g., Searcy & Nowicki, 2005). Our results suggest that both parties do indeed benefit when people signal via environmentalism: environmentalists benefit by broadcasting their cooperative intent (i.e., they are chosen more often and receive higher cooperation within those relationships), and observers benefit by knowing whom to trust. We are not the first to suggest that environmentalism might signal some trait (Griskevicius et al., 2010; van Horen et al., 2018), but we are the first to show why this form of signaling is stable.

5.1. Relation to other research on cooperation

We proposed that environmentalism is a form of cooperation (see also Milinski et al., 2006), and is supported by reputational benefits. Our results strongly support our contention that environmentalism is subject to many of the same principles that apply to other forms of cooperation and reputation, such as sharing food (Smith & Bliege Bird, 2005), religious ritual (Sosis, 2004), or interpersonal generosity (Barclay & Willer, 2007).

Future research should test other principles of signaling cooperative intent, to see if they apply to environmentalism. For example, costly generosity is a better signal of cooperative intent than cheap generosity (e.g., Nelissen, 2008); is the same true about costly versus cheap environmental acts? Many signals do not involve extravagant costs (Barker, Power, Heap, Puurtinen, & Sosis, 2019): cheap rituals can serve as a costly signal of good character if they are repeated often enough (Bliege Bird et al., 2018; Power, 2017); is the same true about cheap environmental behaviours like recycling? Audiences discount others' charitable giving if the actor benefits from that charity (Lin-Healy & Small, 2012); will audiences discount others' environmental behaviour if the actor benefits? Our mathematical model says yes. Altruists are seen as more desirable mates (Arnocky, Piché, Albert, Ouellette, & Barclay, 2017; Barclay, 2010b); are environmentalists seen the same way? When actors are strongly embedded in their communities, they have more opportunities to earn a good reputation from other types of cooperation (Lyle III & Smith, 2014); is the same true of environmentalism? We look forward to studies on environmentalism that test these and other principles of cooperation.

More broadly, our results support the prediction from biological markets theory (Noë & Hammerstein, 1994, 1995) that humans and other organisms will compete over any traits that choosers value, including cooperativeness (Barclay, 2013, 2016). Choosers will value traits that carry statistical information about another's ability, willingness, or availability to confer benefits (e.g., signals of cooperative intent). It is adaptive to choose partners based on such traits, even when they are exaggerated – people's relative ranking on the trait does not change because everyone can exaggerate, so relative values of the trait still carry informational value (Biernaskie et al., 2018). Thus, signals of cooperative intent do not lose their signaling value within biological markets, despite any escalation, because of the focus on *relative* cooperation. Although many theories predict that observation alone will increase cooperation (e.g., indirect reciprocity: Nowak & Sigmund, 2005), biological markets theory (and its renamed variants, e.g., competitive altruism theory: Roberts, 1998; van Vugt, Roberts & Hardy, 2007) is unique in predicting an escalation of these cooperative traits in response to competition to be chosen as a partner. Future research should test for an escalation of other traits that could signal cooperativeness towards potential partners, such as self-sacrifice, loyalty, religiosity, political partisanship, commitment to ingroup ideals ("I'm a level 4 vegan, I don't eat anything that casts a shadow"[2]), and even attacks on a partner's enemies. Biological markets theory predicts escalation – and polarization – of all such traits as a means of competing over partners.

Do our results apply to pro-environmental behaviours that are not readily visible? This question has arisen many times with other cooperative behaviours: if reputational opportunities make people more cooperative, then why do people ever help anonymously? Cooperation researchers have provided many answers (e.g., Barclay & van Vugt, 2015; Burnham & Johnson, 2005; Delton, Krasnow, Cosmides, & Tooby, 2011; Hagen & Hammerstein, 2006; Raihani & Bshary, 2015; West, El Mouden, & Gardner, 2011; Yamagishi, Terai, Kiyonari, Mifune, & Kanazawa, 2007), of which we present five here. First, pro-environmental *sentiment* causes both public and private pro-environmental *behaviour*. As long as the benefits in public outweigh the costs in private, then pro-environmental sentiment will proliferate because it is adaptive on average. Second, people often do find out about many allegedly invisible pro-environmental actions,

[2] "Lisa the Tree Hugger", The Simpsons (2000), season 12 episode 4, written by Matt Selman.

especially if actors mention those acts. For example, energy companies sometimes provide door stickers to customers who use renewable energy; future research should test if these increase renewable energy use in the same way that medals for donating blood increase blood donations (e.g., Lacetera & Macis, 2010). Third, even small environmental actions are hard to hide from spouses, close friends, and others involved in the interaction (e.g., vendors of environmental products). Fourth, error-management (e.g., Haselton & Nettle, 2006) may lead people to err on the side of being pro-environmental even when anonymous, just in case their actions are discovered: it is better to occasionally pay unnecessary costs than to appear hypocritical for being publicly pro-environmental and privately anti-environmental. Fifth, whatever the prevalence of a given pro-environmental act when people are anonymous, such actions can become *more* common if those actions become public and people can earn a good reputation for performing them (e.g., Jacquet, 2015). Therefore, we argue that our findings apply to *all* kinds of pro-environmental behaviour, even those that sometimes appear to be private.

5.2. Comparing different types of reputation for environmentalism

In addition to evaluating our hypothesis of signaling cooperative intent (Barclay, 2012a), we can evaluate alternative reputation-based hypotheses about the function[3] of environmental behaviour, namely indirect reciprocity (Milinski et al., 2006) and signaling wealth (Griskevicius, Cantu & van Vugt, 2012). Table 1 compares predictions from these three hypotheses, as well as two "null theories" (that environmentalism is unrelated to cooperation, or is related to cooperation but not reputation). The theory with the strongest support is that environmentalism is used in partner choice because it signals cooperative intent. We note that the three reputational theories are not mutually exclusive: expensive environmental acts may still function to advertise wealth (Griskevicius, Cantu & van Vugt, 2012), but Table 1 suggests it is probably not the primary function for most kinds of environmentalism, especially low-cost environmentalism (see "mundane help" vs. "extravagant help", Barclay, 2013; Barclay & Reeve, 2012). Indirect reciprocity makes many of the same predictions as signaling cooperative intent, and thus received decent support; in practice these two theories are hard to differentiate (Barclay, 2015). However, only the theory of signaling cooperative intent – the hypothesis we proposed in this paper – successfully predicts all seven major results of our studies.

5.3. Harnessing competition over reputation and limitations thereof

Several authors advocate using reputation and its associated emotions (e.g., guilt, shame) to increase environmental behaviour (e.g., Barclay, 2012a; Griskevicius, Cantu & van Vugt, 2012; Jacquet, 2015). Policy makers have successfully harnessed the power of reputational opportunities to promote cooperation in real-world settings including sustainable energy use (reviewed by Kraft-Todd et al., 2015). Our current results suggest that harnessing *competition* over a good reputation could drive environmentalism even higher (see also Barclay, 2012a; van der Linden, 2015; van Horen et al., 2018). We used three different forms of environmentalism in our four studies, so our results about "competitive environmentalism" could apply to many kinds of socially conscious behaviours such as other charitable donations, purchasing fair trade or "green" items, or reducing one's energy consumption (e.g., Schultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007). Competition could be created by publishing lists (nationally, locally, or intra-organization) that explicitly rank individuals' or organizations' actions to help the environment, or giving relative ratings like "gold" and "silver" to donations or relative sustainability (as some organizations already do). In real life, people choose their social partners (like in the Choice/Knowledge condition), so giving ranked (rather than binary yes/no) information about others' relative environmental efforts would make it easier for the most sustainable individuals to preferentially assort with each other – and cause others to escalate their sustainability as well. In past research, cooperators assort with each other based on behaviours like pre-commitments to charity, and this results in higher cooperation (Brekke et al., 2011; Hauge et al., 2019); we advocate harnessing the same effects with environmental efforts.

We emphasize that would-be social engineers must always beware of the current limitations and unknowns about harnessing reputational opportunities (Barclay, 2012a), lest their attempts be counterproductive. For example, when people receive extrinsic financial incentives for pro-environmental behaviour, it often undermines their intrinsic motivation ("crowding out", e.g., Frey & Jegen, 2001; Gneezy & Rustichini, 2000) or the pro-environmental behaviour may stop when the financial incentives end (van der Linden, 2015). However, this is not true for all extrinsic incentives. More recent evidence shows that non-monetary material incentives (e.g., vouchers) do not undermine intrinsic motivation (Lacetera & Macis, 2010). More importantly, a meta-analysis shows that while financial rewards undermine intrinsic motivation, verbal rewards and positive feedback *increase* intrinsic motivation (Deci, Koestner, & Ryan, 1999). Providing ranked information or relative ratings is a verbal reward rather than a financial incentive – the former allow audiences to confer social rewards, rather than have monetary rewards be imposed top-down. Thus, reputational rewards are unlikely to undermine intrinsic motivation, especially if accompanied by messages of social approval (Schultz et al., 2007). In fact, reputational rewards may even enhance intrinsic motivation (see "crowding in", Weibel, Wiemann, & Osterloh, 2018), and be especially useful for high-cost sustainable behaviour that cannot be maintained by intrinsic motivation alone (Imas, 2014; van der Linden, 2018).

Another potential limitation is that competition over "green-ness" will not necessarily generalize to all populations, nor will the correlation between environmentalism and cooperation. These should be highest when the audience places a high value on the environment, or the actor believes they do (e.g., correlations were higher in Studies 3–4 on a college campus than in Studies 1–2 online with the more general population of Amazon Mechanical Turk). In other words, these effects should be biggest in places where there are norms of environmentalism (Gifford & Nilsson, 2014). As such, effect sizes in other populations may be larger or smaller than those in our experiment, depending on the attributes of the population being tested. The generalizability of any result is ultimately an empirical question (e.g., Henrich, Heine, & Norenzayan, 2010). However, there are at least three reasons to believe our effects will generalize to other populations. First, we found similar effects in both Study 3 and Study 4, despite participants being more supportive of the environment and the Sierra Club in Study 4 (see Supplementary Material, Table S7). Second, although participants' environmentalism was correlated with their absolute donations to the Sierra Club in Studies 3 and 4, their environmentalism was not correlated with how much the experimental conditions affected their donations (Supplementary Material, Table S8; for further discussion on absolute versus relative contributions, see Barker, Barclay, & Reeve, 2012, 2013; Kümmerli, Burton-Chellew, Ross-Gillespie, & West, 2010). In other words, reputational opportunities had a similar effect on all participants' donations, regardless of their own environmentalism. Third,

[3] Please recall that we are discussing *ultimate functions* of environmental behaviours, not *proximate psychological mechanisms* like emotions (e.g., (Scott-Phillips et al., 2011; Tinbergen, 1968); Section 1.3). That is, we are not asking which specific attitudes or emotions cause people to behave environmentally. Instead we are asking why people experience those pro-environmental attitudes or emotions in the first place.

Table 1
Comparison of functional theories about pro-environmentalism as to how well they predict our empirical findings.

	Prediction	Does the Theory Make the Prediction?				
		Indirect Recip.	Signal Wealth/ Status	Signal of Coop.	Other Coop. (No Reputation)	Personal (Unrelated to Coop.)
Predictions with Strong Support	People who act more environmentally are more cooperative towards partners (Pred. 1a)	**Yes**	(Silent)	**Yes**	**Yes**	*No*
	People who do more for the environment have less selfish personalities (Pred. 1b)	**Yes**	(Silent)	**Yes**	**Yes**	*No*
	Cooperative personality mediates the relationship between environmentalism and cooperation, at least partially (Pred. 2)	Possibly	(Silent)	**Yes**	Possibly	*No*
	People act more environmentally when observed (Pred. 3)	**Yes**	**Yes**	**Yes**	*No*	*No (or Silent)*
	People escalate their environmental acts when competing over partners (Pred. 4)	(Silent)	**Yes**	**Yes**	*No*	*No (or Silent)*
	People who act more environmentally are preferentially chosen as partners (Pred. 5)	Possibly	**Yes**	**Yes**	(Silent)	*No (or Silent)*
	People who act more environmentally receive more cooperation (Pred. 6)	**Yes**	Possibly	**Yes**	(Silent)	*No*
Predictions with Partial Support	Environmentalism-cooperation link is weaker when strong incentives exist to appear environmental (tentatively found — Supplementary) (Pred. 7)	(Silent)	(Silent)	**Yes**	*No*	*No (or silent)*
Predictions with Little or No Support	People who value the environment more should reward environmentalists more (Supplementary)	*Yes*	Silent	Silent	Silent	*Yes*
	Wealthier people will do more for the environment (Supplementary)	Silent	*Yes*	Silent	Silent	Silent
	People who possess other desirable traits will do more for the environment – we tested intelligence, attractiveness, & strength (Supplementary)	Silent	*Yes*	Silent	Silent	Silent

For each prediction, we list whether each theory explicitly makes that prediction ("Yes"), does not explicitly make the prediction but it is very compatible with the theory ("Possibly"), neither makes that prediction nor its opposite ("Silent"), or explicitly makes the opposite prediction ("No"). We also note whether the empirical findings of Studies 1–4 strongly support a theory by confirming its explicit predictions (green bolded), rebut a theory by finding the opposite of what it predicted (red italicized), or suggest that a theory is an incomplete explanation at best because it failed to predict an observed finding (yellow bracketed).

environmentalism was correlated with cooperation in all four studies, despite the different populations of Studies 1–2 versus 3–4 and the different measure of environmentalism (a scale of environmental behaviours vs. monetary donations vs. attitudes towards the Sierra Club). These all speak towards the generalizability of our findings.

The environmental movement is typically associated with left-wing politics, at least in North America. MTurk is ideologically diverse, and we found a link between environmentalism and cooperation on MTurk. Would we also find our results in much more conservative populations? Possibly, if they recognized that environmentalists were more likely to cooperate — everyone benefits from trusting a good person. This recognition is especially likely if the environmental cause were something that conservatives cared about and benefited from. For example, the non-profit organization Ducks Unlimited protects wetlands for duck hunting, and some conservative groups protect ranches and farms from industrial contamination. Republican president Richard Nixon responded to toxic pollution by creating the Environmental Protection Agency. This suggests that environmentalism – and responses to it – are not inherently partisan. We suggest that conservatives would respond just as highly to conservative environmental causes as liberals respond to liberal causes, and all of our effects would apply to both groups (especially if the environmental cause was framed with moral concerns that conservatives care about, see (Feinberg & Willer, 2013)). However, this is an empirical question — future research should test it.

Ultimately, for environmentalists to benefit from having a cooperative reputation, their partners must be aware that people benefit when others protect the environment. The more people talk about, promote, or publish about environmental sustainability, the more opportunities there are for reputational pressures to promote environmental sustainability. The more embedded someone is within their social network, the more opportunities they have to benefit from a pro-environmental reputation. Costs also matter: environmentalism is more likely to be supported by reputational pressures when the link with cooperation is clear, i.e., when the pro-environmental actions are personally costly and the benefits to others are obvious. In such cases, audiences could come to value environmentalism as they learn that it is a useful indicator of others' cooperative intent or long-term thinking. This can initiate a ratchet-like process that drives environmentalism to higher and higher levels (Barclay, 2011). Despite this escalation, environmentalism will not lose its association with intrinsic cooperativeness, because at equilibrium, each person's *absolute* environmentalism is higher but their *ranking* in environmentalism remains the same relative to others (see Biernaskie et al., 2018). Finally, education about environmental responsibility works synergistically with reputation: education helps create reputational pressures by clarifying environmentalism's link with cooperation, but reputational pressures give people a reason to care (Barclay, 2012a).

CRediT authorship contribution statement

PB conceived of the idea, co-designed the studies, acquired SSHRC funding, created the mathematical models, oversaw the running of Studies 1-2, conducted the statistical analyses, created the figures and tables, and wrote the manuscript first draft and supplementary material. JLB co-designed the studies, conducted Studies 3-4, designed the protocol and spreadsheets to conduct Studies 3-4, and contributed to writing subsequent manuscript drafts and supplementary material.

Further declarations

This manuscript has not been previously published and is not under consideration elsewhere, including the internet. We confirm that we have reported all independent and dependent variables and have excluded no observations. We obtained informed consent from participants after explaining the nature and possible consequences of the study. Our methods were approved by the Institutional Review Board at the University of Arizona and the Research Ethics Board at the University of Guelph.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Sara Kafashan and Jillian O'Connor for programming and posting Studies 1-2, Autri Hafezi, Raymond Kwong and Kellen Fortier for help conducting Studies 3-4, Maxwell Burton-Chellew and Christopher von Rueden for comments on the manuscript, and Sander van der Linden and four anonymous reviewers for peer-reviewed comments and editorial guidance. This work was funded by the National Institute of Health of the USA (grant 5K12GM000708-13), the Social Sciences and Humanities Research Council of Canada (grant 430287) and the Aarhus Institute of Advanced Studies — Marie Curie COFUND Fellowship.

Supplementary material

https://osf.io/3agx5/?view_only=4d47e8cccfe24afab6f6eeef0b350c06.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jenvp.2020.101441.

References

André, J.-B. (2010). The evolution of reciprocity: social types or social incentives? *American Naturalist*, *175*, 197–210.

Andreoni, J., & Bernheim, B. D. (2009). Social image and the 50-50 norm: a theoretical and experimental analysis of audience effects. *Econometrica*, *77*, 1607–1636.

Arnocky, S., Piché, T., Albert, G., Ouellette, D., & Barclay, P. (2017). Altruism predicts mating success in humans. *British Journal of Psychology*, *108*, 416–435.

Axelrod, R. (1984). *The evolution of cooperation*. New York, NY: Basic Books.

Balliet, D., Parks, C., & Joireman, J. (2009). Social value orientation and cooperation in social dilemmas: a meta-analysis. *Group Processes and Intergroup Relations*, *12*(4), 533–547.

Barclay, P. (2004). Trustworthiness and competitive altruism can also solve the tragedy of the commons. *Evolution & Human Behavior*, *25*, 209–220.

Barclay, P. (2010a). *Reputation and the evolution of generous behavior*. Hauppauge, NY: Nova Science Publishers.

Barclay, P. (2010b). Altruism as a courtship display: Some effects of third-party generosity on audience perceptions. *British Journal of Psychology*, *101*, 123–135.

Barclay, P. (2011). Competitive helping increases with the size of biological markets and invades defection. *Journal of Theoretical Biology*, *281*, 47–55.

Barclay, P. (2012a). Harnessing the power of reputation: strengths and limits for promoting cooperative behaviours. *Evolutionary Psychology*, *10*(5), 868–883.

Barclay, P. (2012b). Proximate and ultimate causes of strong reciprocity and punishment. *Behavioral and Brain Sciences*, *35*(1), 16–17.

Barclay, P. (2013). Strategies for cooperation in biological markets, especially for humans. *Evolution & Human Behavior*, *34*(3), 164–175.

Barclay, P. (2015). Reputation. In D. Buss (Ed.), *Handbook of evolutionary psychology* (2nd ed.). (pp. 810–828). Hoboken, NJ: J. Wiley & Sons.

Barclay, P. (2016). Biological markets and the effects of partner choice on cooperation and friendship. *Current Opinion in Psychology*, *7*, 33–38.

Barclay, P. (2020). Reputation and reciprocity: proximate or ultimate causes of human cooperation? (submitted for publication).

Barclay, P., & Reeve, H. K. (2012). The varying relationship between helping and individual quality. *Behavioral Ecology*, *23*(4), 693–698.

Barclay, P., & van Vugt, M. (2015). The evolutionary psychology of human prosociality: adaptations, mistakes, and byproducts. In D. Schroeder, & W. Graziano (Eds.), *Oxford handbook of prosocial behavior* (pp. 37–60). Oxford, UK: Oxford University Press.

Barclay, P., & Willer, R. (2007). Partner choice creates competitive altruism in humans. *Proceedings of the Royal Society of London Series B*, *274*, 749–753.
Barker, J., Barclay, P., & Reeve, H. K. (2012). Within-group competition reduces cooperation and payoffs in human groups. *Behavioral Ecology*, *23*(4), 735–741.
Barker, J., Barclay, P., & Reeve, H. K. (2013). Competition over personal resources favors contributions to shared resources in human groups. *PLoS One*, *8*(3), Article e58826.
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*, *28*(28), 86–99.
Biernaskie, J. M., Perry, J. C., & Grafen, A. (2018). A general model of biological signals, from cues to handicaps. *Evolution Letters*, *2*(3), 201–209.
Bliege Bird, R., Ready, E., & Power, E. (2018). The social significance of subtle signals. *Nature Human Behaviour*, *2*, 452–457.
Bolle, F. (2001). Why to buy your darling flowers: on cooperation and exploitation. *Theory and Decision*, *50*, 1–28.
Bradley, A., Lawrence, C., & Ferguson, E. (2018). Does observability affect prosociality? *Proceedings of the Royal Society of London B*, *285*, Article 20180116.
Brekke, K. A., Hauge, K. E., Lind, J. T., & Nyborg, K. (2011). Playing with the good guys. A public goods game with endogenous group formation. *Journal of Public Economics*, *95*, 1111–1118.
Burnham, T. C., & Johnson, D. D. P. (2005). The biological and evolutionary logic of human cooperation. *Analyse & Kritik*, *27*, 113–135.
Cameron, L. D., Brown, P. M., & Chapman, J. G. (1998). Social value orientations and decisions to take proenvironmental action. *Journal of Applied Social Psychology*, *28*, 675–697.
Casey, P. L., & Scott, K. (2006). Environmental concern and behaviour in an Australian sample within an ecocentric-anthropocentric framework. *Australian Journal of Psychology*, *58*, 57–67.
Curry, O. S., Price, M. E., & Price, J. G. (2008). Patience is a virtue: cooperative people have lower discount rates. *Personality and Individual Differences*, *44*, 778–783.
de Groot, J. I. M., & Steg, L. (2008). Value orientations to explain beliefs related to environmental significant behavior. *Environment and Behavior*, *40*(3), 330–354.
Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analystic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*, 627–668.
Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 13335–13340.
Ekström, M. (2011). Do watching eyes affect charitable giving? Evidence from a field experiment. *Experimental Economics*, *15*, 530–546.
Ernest-Jones, M., Nettle, D., & Bateson, M. (2011). Effects of eye images on everyday cooperative behavior: a field experiment. *Evolution and Human Behaviour*, *32*, 172–178.
Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behaviour*, *25*, 63–87.
Fehrler, S., & Przepiorka, W. (2013). Charitable giving as a signal of trustworthiness: disentangling the signaling benefits of altruistic acts. *Evolution and Human Behaviour*, *34*, 139–145.
Feinberg, M., & Willer, R. (2013). The moral roots of environmental attitudes. *Psychological Science*, *24*, 56–62.
Feinberg, M., Willer, R., & Schultz, M. (2014). Gossip and ostracism promote cooperation in groups. *Psychological Science*, *25*, 656–664.
Francey, D., & Bergmüller, R. (2012). Images of eyes enhance investments in a real-life public good. *PLoS One*, *7*, 1–7.
Frank, R. H. (2004). What price the moral high ground? In *Ethical dilemmas in competitive environments*. Princeton, NJ: Princeton University Press.
Frey, B. S., & Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys*, *15*, 589–611.
Gärling, T., Fujii, S., Gärling, A., & Jakobsson, C. (2003). Moderating effects of social value orientation on determinants of proenvironmental behavior intention. *Journal of Environmental Psychology*, *23*, 1–9.
Gifford, R., & Nilsson, A. (2014). Personal and social factors that influence pro-environmental concern and behaviour: a review. *International Journal of Social Psychology*, *49*(3), 141–157.
Gneezy, U., & Rustichini, A. (2000). Pay enough or don't pay at all. *Quarterly Journal of Economics*, *115*, 791–810.
Griskevicius, V., Cantu, S., & van Vugt, M. (2012). The evolutionary bases for sustainable behavior. *Journal of Public Policy and Marketing*, *31*, 115–128.
Griskevicius, V., Tybur, J. M., Ackerman, J. M., Delton, A. W., Robertson, T. E., & White, A. E. (2012). The financial consequences of too many men: sex ratio effects on saving, borrowing, and spending. *Journal of Personality and Social Psychology*, *102*(1), 69–80.
Griskevicius, V., Tybur, J. M., & Van den Bergh, B. (2010). Going green to be seen: status, reputation, and conspicuous conservation. *Journal of Personality and Social Psychology*, *98*(3), 392–404.
Groot, J. I. M., & Steg, L. (2008). Value orientations to explain beliefs related to environmental significant behavior: How to measure egoistic, altruistic, and biospheric value orientations. *Environment and Behavior*, *40*(3), 330–354.
Hagen, E. H., & Hammerstein, P. (2006). Game theory and human evolution: a critique of some recent interpretations of experimental games. *Theoretical Population Biology*, *69*, 339–348.
Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: an integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review*, *10*, 47–66.
Hauge, K. E., Brekke, K. A., Nyborg, K., & Lind, J. T. (2019). Sustaining cooperation through self-sorting: the good, the bad, and the conditional. *Proceedings of the National Academy of Science of the USA*, *116*(12), 5299–5304.
Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*, 61–135.
Imas, A. (2014). Working for the warm glow: on the benefits and limits of prosocial incentives. *Journal of Public Economics*, *114*, 14–18.
Jacquet, J. (2015). *Is Shame Necessary? New Uses for an Old Tool*. New York, NY: Pantheon.
Jarmasz, J., & Hollands, J. G. (2009). Confidence intervals in repeated-measures designs: the number of observations principle. *Canadian Journal of Experimental Psychology*, *63*, 124–138.
Joireman, J. A., van Lange, P. A. M., & van Vugt, M. (2004). Who cares about the environmental impact of cars? Those with an eye towards the future. *Environment and Behavior*, *36*(2), 187–206.
Joireman, J. A., Lasane, T. P., Bennett, J., Richards, D., & Solaimani, S. (2001). Integrating social value orientation and the consideration of future consequences within the extended norm activation model of proenvironmental behaviour. *British Journal of Social Psychology*, *40*, 133–155.
Jordan, J. J., Hoffman, M., Bloom, P., & Rand, . D. G. (2016). Third-party punishment as a costly signal of trustworthiness. *Nature*, *530*, 473–476.
Keller, J., & Pfattheicher, S. (2011). Vigilant self-regulation, cues of being watched and cooperativeness. *European Journal of Personality*, *25*, 363–372.
Keser, C. (2003). Experimental games for the design of reputation management systems. *IBM Systems Journal*, *42*, 498–506.
Kormos, C., & Gifford, R. (2014). The validity of self-report measures of proenvironmental behavior: a meta-analytic review. *Journal of Environmental Psychology*, *40*, 359–371.
Kraft-Todd, G., Yoeli, E., Bhanot, S., & Rand, D. (2015). Promoting cooperation in the field. *Current Opinion in Behavioral Sciences*, *3*, 96–101.
Kümmerli, R., Burton-Chellew, M. N., Ross-Gillespie, A., & West, S. A. (2010). Resistance to extreme strategies, rather than prosocial preferences, can explain human cooperation in public goods games. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 10125–10130.
Lacetera, N., & Macis, M. (2010). Do all material incentives for pro-social activities backfire? The response to cash and non-cash incentives for blood donations. *Journal of Economic Psychology*, *31*, 738–748.
Lin-Healy, F., & Small, D. A. (2012). Cheapened altruism: discounting personally affected prosocial actors. *Organizational Behavior and Human Decision Processes*, *117*, 269–274.
Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subjects designs. *Psychonomic Bulletin and Review*, *1*, 476–490.
Lyle III, H. F., & Smith, E. A. (2014). The reputational and social network benefits of prosociality in an andean community. *Proceedings of the National Academy of Science of the USA*, *111–113*, 4820–4825.
Milinski, M., Semmann, D., & Krambeck, H.-J. (2002). Reputation helps solve the tragedy of the commons. *Nature*, *415*, 424–426.
Milinski, M., Semmann, D., Krambeck, H.-J., & Marotzke, J. (2006). Stabilizing the Earth's climate is not a losing game: supporting evidence from public goods experiments. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 394–3998.
Murphy, R. A., & Ackermann, K. A. (2014). Social value orientation: theoretical and measurement issues in the study of social preferences. *Personality and Social Psychology Review*, *18*(1), 13–41.
Murphy, R. O., Ackermann, K. A., & Handgraaf, M. J. J. (2011). Measuring social value orientation. *Judgment and Decision Making*, *6*(8), 771–781.
Muthukrishna, M., & Henrich, J. (2019). A problem in theory. *Nature Human Behaviour*, *3*, 221–229.
Nelissen, R. (2008). The price you pay: cost-dependent reputation effects of altruistic punishment. *Evolution & Human Behavior*, *29*(4), 242–248.
Noë, R., & Hammerstein, P. (1994). Biological markets: supply and demand determine the effect of partner choice in cooperation, mutualism and mating. *Behavioral Ecology & Sociobiology*, *35*, 1–11.
Noë, R., & Hammerstein, P. (1995). Biological markets. *Trends in Ecology and Evolution*, *10*, 336–339.
Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, *437*, 1291–1298.
Ohtsubo, Y., & Watanabe, E. (2009). Do sincere apologies need to be costly? Test of a costly signaling model of apology. *Evolution and Human Behaviour*, *30*, 114–123.
Pfeiffer, T., Tran, L., Krumme, C., & Rand, D. G. (2012). The value of reputation. *Journal of the Royal Society Interface*, *9*, 2791–2797.
Pieters, R. (2017). Meaningful mediation analysis: plausible causal inference and informative communication. *Journal of Consumer Research*, *44*(3), 692–716.
Power, E. (2017). Testing the signaling theory of religion. *Evolution and Human Behaviour*, *38*, 82–91.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879–891.
Raihani, N. J., & Bshary, R. (2015). Why humans might help strangers. *Frontiers in Behavioral Neuroscience*, *9*, 1–11.
Raihani, N. J., & Smith, S. (2015). Competitive helping in online giving. *Current Biology*, *25*(9), 1183–1186.
Rand, D. G. (2012). The promise of Mechanical Turk: how online labor markets can help theorists run behavioral experiments. *Journal of Theoretical Biology*, *299*, 172–179.
Roberts, G. (1998). Competitive altruism: from reciprocity to the handicap principle. *Proceedings: Biological Sciences*, *265*, 427–431.
Roberts, G., & Renwick, J. S. (2003). The development of cooperative relationships: an experiment. *Proceedings: Biological Sciences*, *270*, 2279–2283.
Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, *18*(5), 429–434.
Scott-Phillips, T. C., Dickins, T. E., & West, S. A. (2011). Evolutionary theory and the ultimate-proximate distinction in the human behavioral sciences. *Perspectives on Psychological Science*, *6*(1), 38–47.
Searcy, W. A., & Nowicki, S. (2005). *The evolution of animal communication: Reliability and deception in signaling systems*. Princeton, NJ: Princeton University Press.
Servedio, M. R., Brandvain, Y., Dhole, S., Fitzpatrick, C., Goldberg, E., Stern, C., Van Cleve, J., & Yeh, J. (2014). Not just a theory – the utility of mathematical models in evolutionary biology. *PLoS Biology*, *12*(12), Article e1002107.
Sexton, S. E., & Sexton, A. L. (2014). Conspicuous conservation: the prius halo and willingness to pay for environmental bona fides. *Journal of Environmental Economics and Management*, *67*, 303–317.
Smith, E. A., & Bliege Bird, R. (2005). Costly signalling and cooperative behaviour. In H. Gintis, S. Bowles, R. Boyd, & E. Fehr (Eds.), *Moral sentiment and material interests: The foundations of cooperation in economic life* (pp. 115–148). Cambridge, MA: MIT Press.
Snyder, M., & Gangestad, S. (1986). On the nature of self-monitoring, matters of assessment, matters of validity. *Journal of Personality and Social Psychology*, *51*(1), 125–139.
Sosis, R. (2004). The adaptive value of religious ritual. *American Scientist*, *92*, 166–172.
Steg, L., Bolderdijk, J. W., Keizer, K., & Perlaviciute, G. (2014). An integrated framework for encouraging pro-environmental behaviour: the role of values, situational factors and goals. *Journal of Environmental Psychology*, *38*, 104–115.
Steg, L., & Vlek, C. (2009). Encouraging pro-environmental behaviour: an integrative review and research agenda. *Journal of Environmental Psychology*, *29*, 309–317.
Sussman, R., Lavalle, L. F., & Gifford, R. (2016). Pro-environmental values matter in competitive but not cooperative commons dilemmas. *The Journal of Social Psychology*, *156*(1), 43–55.
Sylwester, K., & Roberts, G. (2010). Cooperators benefit through reputation-based partner choice in economic games. *Biology Letters*, *6*, 659–662.
Tinbergen, N. (1968). On war and peace in animals and man. *Science*, *160*, 1411–1418.
Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323.
van der Linden, S. (2015). Intrinsic motivation and pro-environmental behaviour. *Nature Climate Change*, *5*, 612–613.
van der Linden, S. (2018). Warm glow is associated with low- but not high-cost sustainable behaviour. *Nature Sustainability*, *1*, 28–30.
van Horen, F., van der Wal, A., & Grinstein, A. (2018). Green, greener, greenest: can competition increase sustainable behavior? *Journal of Environmental Psychology*, *59*, 16–25.
van Vugt, M., Meertens, R. M., & van Lange, P. A. M. (1995). Car versus public transportation? The role of social value orientations in a real-life social dilemma. *Journal of Applied Social Psychology*, *25*, 258–278.
van Vugt, M., Roberts, G., & Hardy, C. (2007). Competitive altruism: a theory of reputation-based cooperation in groups. In R. Dunbar, & L. Barrett (Eds.), *Oxford handbook of evolutionary psychology* (pp. 531–540). Oxford, UK: Oxford University Press.
Weibel, A., Wiemann, & Osterloh, M. (2018). A behavioral economics perspective on the overjustification effect: crowding-in and crowding-out of intrinsic motivation. In M. Gagné (Ed.), *The Oxford handbook of work engagement, motivation, and self-determination theory* (pp. 72–84). Oxford, UK: Oxford University Press.
West, S. A., El Mouden, C., & Gardner, A. (2011). Sixteen misconceptions about the evolution of cooperation in humans. *Evolution and Human Behaviour*, *32*, 231–262.
Wu, J., Balliet, D., & Van Lange, P. A. M. (2016). Gossip versus punishment: the efficiency of reputation to promote and maintain cooperation. *Scientific Reports*, *6*(23919).
Yamagishi, T., Terai, S., Kiyonari, T., Mifune, N., & Kanazawa, S. (2007). The social exchange heuristic: managing errors in social exchange. *Rationality and Society*, *19*(3), 259–291.