



How to Promote Cooperation for the Well-Being of Individuals and Societies

Valerio Capraro^(✉)

Middlesex University, London NW44BT, UK
caprarovalerio@gmail.com

Abstract. The most successful human societies are those that have found better ways to promote cooperative behaviour. Yet, cooperation is individually costly and, therefore, it often breaks down, leading to enormous social costs. In this article, I review the literature on the mechanisms and interventions that are known to promote cooperative behaviour in social dilemmas. In iterated or non-anonymous interactions, I focus on the five rules of cooperation, as well as on structural changes, involving the cost or the benefit of cooperation, or the size of the interacting group. In one-shot and anonymous interactions, I focus on the role of internalised social heuristics as well as moral preferences for doing the right thing. For each account, I summarize the available experimental evidence. I hope that this review can be helpful for social scientists working on cooperation and for leaders and policy makers who aim at promoting social cooperation or teamwork.

Keywords: Cooperation · Prisoner's Dilemma · Public goods game · Social heuristics · Moral suasion

1 Introduction

Cooperative behaviour is defined as paying a cost to give a greater benefit to one or more other people. Since the benefit is greater than the cost, cooperation increases the total payoff of the group made of all people involved in the interaction. For this reason, cooperative behaviour is considered by social scientists to be one of the key ingredients for a successful society (Boyd et al. 2003; Fehr and Gächter 2002; Fehr and Fischbacher 2003; Tomasello et al. 2005; Nowak, 2006; Rand and Nowak 2013; Perc et al. 2017). Moreover, while individual cooperation increases social well-being, collective cooperation not only increases social well-being, but also increases personal well-being: each individual within a society made of cooperators is better off than each individual within a society made of defectors.

Yet, since cooperation is individually costly, it often breaks down. So, one of the most important research programs across social sciences seeks to find ways to promote and sustain cooperative behaviour. In this article, I will review the literature on this topic.

Obviously, this field of research is enormous, and, in the limited space of these pages, I can only scratch its surface. To try to counterbalance this, I will include several references, where the interested reader can find more detailed information. I hope that this review can be a useful starting point for social scientists, policy makers and leaders who are interested in how to promote cooperative behaviour.

2 Models of Cooperation

Cooperation is formally studied through social dilemmas. These are strategic interactions in which $N > I$ individuals get to decide between two or more actions. Among these actions, there is one that benefits the group and one that benefits the individual. This tension between self-interest and collective interest is what defines a *social dilemma*. Moving down from general to particular, social scientists have defined different social dilemmas meant to conceptualise cooperative behaviour in different prototypical circumstances. In this article, I will focus on the two most-studied social dilemmas: the prisoner's dilemma and the public goods game.

2.1 Cooperation Between Two Agents: The Prisoner's Dilemma

In the prisoner's dilemma there are two players, each of whom has two available strategies, cooperate or defect. Cooperators pay a cost c to give a greater benefit b to the other player. Defectors pay no cost and generate no benefit.

The conflict between individual and collective interest descends from the assumption $b > c$. Indeed, if both individuals cooperate, they each pay the cost of cooperation but also enjoy its benefit; therefore, they each get $b - c$. However, each individual has an incentive to deviate from cooperation to get the full payoff b , obtained by saving their own cost of cooperation, while keeping the benefit of the other's cooperative act. However, if both individuals reason this way, they both end up with 0, which is smaller than $b - c$, the payoff that they would have gotten if they had both resisted the temptation to defect.

There are also other ways to formalise cooperative behaviour between two agents, including the traveller's dilemma (Basu 1994), the Bertrand competition (Bertrand 1883), and the centipede game (Binmore 1987). While I acknowledge the importance of these models, in this article I will focus on the prisoner's dilemma (for two individuals), because the theoretical mechanisms and the experimental regularities discussed below have been developed and tested mainly using the prisoner's dilemma, but they are expected to work, in a similar fashion, in the other 2-player social dilemmas.

2.2 Cooperation Among N Agents: The Public Goods Game

The most popular N -player social dilemma is the public goods game, which is meant to conceptualise situations in which a group of individuals get to decide how much to contribute to a common project.

Formally, in the public goods game, each of N individuals gets to decide how much, if any, of their initial endowment e they want to contribute to the public good. Let c_i be the contribution of individual i , the total amount contributed by all individuals is then $c_1 + \dots + c_N$. Let a be the "marginal return of cooperation", the payoff of individual i is defined as $e - c_i + a(c_1 + \dots + c_N)$, that is, i receives the portion of the endowment that she decided to keep plus a proportion a of the common good. The marginal return of cooperation a is assumed to be greater than $1/N$ and smaller than 1 . This assumption guarantees that the public goods game is a social dilemma: the collective interest would be maximised if all individuals contribute the whole endowment, but each individual has an incentive to deviate and keep the whole endowment.

There are also other ways to conceptualise cooperation among N individuals, as the N -player prisoner's dilemma and the piecewise linear-then-constant public goods game, among others (threshold public goods game, resource dilemma, volunteer's dilemma, etc.). Although in this article I will not focus on these games, I think it is worth defining them because this allows to shed light on the different effects that group size can have on cooperative behaviour depending on the social dilemma.

In the public goods game as defined above, there is the underlying assumption that the individual return for full cooperation increases linearly with the number of individuals, that is, if all individuals cooperate, then each of them gets $e*a*N$, which increases linearly with N . In some practical contexts, however, this assumption is unrealistic. For studying these situations, one can consider social dilemmas where the assumption of linearity of the relationship between group size and individual return for full cooperation is replaced with other assumptions. Here, I discuss two prototypical cases. One is the N -player prisoner's dilemma. Over the years, several definitions of this game have been proposed at various levels of generality (e.g., Hamburger 1973; Carroll 1988). Here, I define an N -player prisoner's dilemma to be any N -player social dilemma where the individual return for full cooperation is constant with the number of players. This game is useful to formalise situations in which the individual benefit of cooperation does not depend on the number of players, but, still, one needs all players to cooperate (Yao and Darwen 1994; Grujić et al. 2012; Barcelo and Capraro 2015). Another practically relevant N -player social dilemma is the piecewise linear-then-constant public goods game, where the return of cooperation increases linearly until a certain group size N_0 , and then becomes constant. This conceptualises situations in which the production of the public good reaches a *plateau* due to natural limits in the production (Yang et al. 2013; Capraro and Barcelo 2015).

3 Iterated or Non-anonymous Interactions

Most of our everyday interactions are repeated or with people who are not completely anonymous. For example, we may interact with a friend of a friend, or with a company that has been recommended to us. In these cases, cooperation can evolve even among self-interested agents, according to five fundamental rules that have been summarised by Nowak (2006) in the case of the prisoner's dilemma. I review these rules below. For each rule, I will also review the experimental evidence. Most of the experimental work is taken from Rand and Nowak (2013)'s review, which I recommend for further details.

3.1 Kin Selection

Kin selection allows to explain cooperation between relatives. The general assumption of the theory is that, if r is the probability of sharing a gene, then an individual does not only receive its payoff, but also a proportion r of the others' payoff. Applying this to the prisoner's dilemma, it follows that, if $r*b > c$, then it becomes individually optimal to cooperate. This rule takes the name of Hamilton's rule, from the pioneering work of biologist William D. Hamilton (e.g., Hamilton 1964). The experimental evidence in support of this rule is, however, scarce, mainly because it is difficult to isolate the effect

of genetic relatedness from other components that are usually associated with genetic relatedness, such as long-term relationships and the possibility of future interactions. Despite these technical difficulties, Madsen et al. (2007) were able to analyse data from two different cultures while controlling for three potential sources of confound, generational effects, sexual attraction, and reciprocity. In doing so, they found that people behaved in accordance with Hamilton's rule.

3.2 Direct Reciprocity

Direct reciprocity permits to explain the evolution of cooperation in the context of repeated interactions between the same individuals. In this case, I can cooperate with you today, to receive the benefit of your future cooperation tomorrow. However, this leads to a cooperative equilibrium only when the probability of another encounter is large enough, so that the future benefit of cooperation, discounted by the probability of another encounter, is greater than the present cost of cooperation. Formally, let w be this probability of another encounter, one needs $w*b > c$. Experimental studies have shown that, indeed, the rate of cooperation in indefinitely iterated social dilemmas increases when the probability of future encounters increases (Roth and Murnighan 1978; Murnighan and Roth 1983; Duffy and Ochs 2009; Dal Bó and Fréchet 2011; Fudenberg et al. 2012).

3.3 Indirect Reciprocity

In case of repeated interactions with rematching after each round, people may become more likely to cooperate when they have information about the others' reputation. In its simplest form, reputation simply coincides with the behaviour in the past interaction. In this case, people can selectively cooperate with those who have cooperated in the previous round. Anticipating this, people may become more inclined to cooperate from the first round. Several experiments have indeed shown that people tend to cooperate with people who have cooperated in the past and that the presence of a reputational mechanism can promote and sustain cooperation (Bolton et al. 2005; Milinski et al. 2006; Rockenbach and Milinski 2006; Seinen and Schram 2006; Rand et al. 2009; Pfeiffer et al. 2012). The fact that people assign value to knowing others' behaviour is shown also by the fact that people invest a lot of time in acquiring information about the behaviour of others (Dunbar et al. 1997; Sommerfeld et al. 2007). Indeed, it can be shown that cooperation can be supported by indirect reciprocity only if the probability of knowing someone's reputation is greater than c/b (Nowak and Sigmund 1998).

3.4 Network Reciprocity

Most human interactions are not random, but structured. Network reciprocity allows to explain the evolution of cooperation on graphs, where nodes represent actors and edges represent interactions between actors. The idea is that, if interactions are structured, then clusters of cooperators can protect themselves from the invasion of defectors. This however requires that the ratio b/c is large enough. A simple rule that works on many

graphs and with several strategy updating mechanisms is $b/c > k$, where k is the average degree of the graph (Ohtsuki et al. 2006). Rand et al. (2014a) showed experimentally that cooperation indeed can evolve in graphs satisfying this rule. Instead, if this rule is not satisfied, the rate of cooperation in structured populations is typically the same as in well-mixed populations (Grujić et al. 2010; Traulsen et al. 2010; Suri and Watts 2011; Gracia-Lázaro et al. 2012; Grujić et al. 2012a). Some work also explored the evolution of cooperation on dynamic networks, where people can break old links and create new ones after each interaction. It has been found that people tend to break links with defectors and create links with cooperators, and this leads to an additional increase in the rate of cooperation, compared to static networks, both in mathematical models (Bilancini & Boncinelli, 2009; Bilancini et al. 2018) and in economic experiments (Fehl et al. 2011; Rand et al. 2011; Wang et al. 2012).

3.5 Group Selection

When there is group selection, that is, competition between groups, groups of cooperators might outperform groups of defectors, leading to the evolution of cooperation (Richerson et al. 2016). A mathematically simple necessary condition for the evolution of cooperation can be found assuming rare group splitting and weak selection: let n be the maximum group size and m the number of groups, cooperation may evolve only if $b/c > 1 + n/m$ (Traulsen and Nowak 2006). To the best of my knowledge, this specific formula has not been tested experimentally. However, there is abundant evidence that the presence of intergroup competition can lead to the evolution of intragroup cooperation (Erev et al. 1993; Gunnthorsdottir and Rapoport 2006; Puurtinen and Mappes 2009), even when there is no monetary prize associated with winning the competition (Tan and Bolle 2007; Böhm and Rockenbach 2013).

3.6 Cost, Benefit, and Group Size

Beyond the five mechanisms above, there are also structural changes in the social dilemmas that might promote the evolution of cooperation, by facilitating the application of one of the five rules of cooperation. Two of these structural changes are straightforward. In fact, the five mathematical conditions described above imply that, when b increases or c decreases, the evolution of cooperation becomes easier, in the sense that the set of the values of the other parameters for which cooperation can evolve grows larger. Experimentally, the fact that cooperative behaviour in iterated prisoner's dilemmas depends positively on b and negatively on c was already observed in the early book of Rapoport and Chammah (1965). Similar findings have been reported also in the iterated public goods game, where an increase in the marginal return of cooperation a is typically associated with an increase in cooperative behaviour, both with partner- and with random-rematching (Gunnthorsdottir et al. 2007).

For social dilemmas with $N > 2$ players there is another parameter that may affect cooperative behaviour, group size. However, the effect of group size on cooperation depends on the type of social dilemma: in the iterated public good game, larger groups tend to cooperate more (Isaac et al. 1994), whereas, in the iterated prisoner's dilemma, larger groups tend to cooperate less (Grujić et al. 2012b). The intuition behind this result

is that, in the prisoner's dilemma, the individual return for full cooperation is constant as the group size increases, so it becomes more and more difficult to get the same payoff, and this may work as an incentive to defect; on the other hand, in the public goods game, the individual return for full cooperation increases linearly with the group size, and this might incentivise people to cooperate, despite a potentially larger absolute number of defectors.¹

3.7 Punishment and Reward

Numerous studies using iterated social dilemmas have shown that the presence of punishment or reward tend to increase cooperative behaviour (Yamagishi et al. 1986; Ostrom et al. 1992; Fehr & Gächter, 2000; Rand et al. 2009), despite the presence of occasional anti-social punishers, people who punish cooperative behaviour (Herrmann et al. 2008). In other words, if people interact knowing that they might be punished or rewarded for their behaviour, they tend to cooperate more. This happens because, on average, people tend to punish defectors and reward cooperators. This finding can be interpreted as a form of reciprocity (direct or indirect, depending on whether the punisher/rewarder is affected by the choice of the defector/cooperator) and therefore it could be seen as a way in which one can apply two of the five rules of cooperation in reality; indeed, institutional punishment of defectors is perhaps the oldest-known way to promote cooperative behaviour within a society.

4 One-Shot and Anonymous Interactions

In one-shot and anonymous games, the standard theory of rational, payoff-maximising behaviour predicts that people never cooperate. However, behavioural experiments have repeatedly shown that some people do cooperate even in these contexts (Rapoport and Chammah 1965). Usually, the structural changes in the strategic interaction that promote cooperation in iterated games (while maintaining the anonymity of the interactions) promote cooperation also in one-shot games. Specifically, decreasing the cost of cooperation (Engel and Zhurakhovska 2016) or increasing its benefit (Capraro et al. 2014) do promote cooperative behaviour; increasing the size of the group promotes cooperative behaviour in the public goods games, but reduces it in the N -player prisoner's dilemma (Barcelo and Capraro 2015). In one-shot anonymous games and in iterated games with random-matching, there is also some research on the effect of group size on cooperation in the piecewise linear-then-constant public goods game, but the results are mixed: one study found an inverted-U relationship, such that intermediate size groups were the most cooperative (Capraro & Barcelo 2015), but a subsequent study failed to replicate this finding and found a positive effect of group size on cooperation (Pereda et al. 2019).

The presence of punishment or reward increases cooperation also in one-shot games (Capraro et al. 2016; Capraro and Barcelo 2021a). Moreover, cues that suggest that the

¹ This suggests that there might be intermediate cases in which the individual return of full cooperation increases too slowly with the group size, leading to a null or even a negative effect of group size on cooperation. For example, it could be interesting to study the relationship between group size and cooperation in a logarithmic public goods game.

interaction may not be anonymous or one-shot can increase cooperative behaviour. For example, information about the other participants' behaviour can increase cooperation, via conditional cooperation (Fischbacher et al. 2001; Kocher et al. 2008).

However, this does not explain why people cooperate in one-shot and anonymous social dilemmas. In this section, I review the main frameworks that have been proposed in the last decade.

4.1 Social Heuristics

One prominent account contends that people internalise strategies that are useful in their everyday life and use them as heuristics when they happen to interact in novel situations. Specifically, most real-life interactions are not one-shot and anonymous, but they are repeated or non-anonymous, they happen with friends or colleagues, or with individuals about whom we have information regarding their past behaviour. In these contexts, people may learn that cooperative behaviour pays off in the long run, especially when its cost is low, or its benefit is high (or in groups of special size). People might then learn and internalise these heuristics and apply them in one-shot and anonymous games. This framework takes the name of Social Heuristics Hypothesis (Rand et al. 2014b).

Over the last decade, scholars have sought to experimentally test this framework. The idea behind the experimental approach is the following: if cooperation in one-shot and anonymous games is driven by heuristics, then experimental manipulations aimed at increasing reliance on heuristics should promote cooperative behaviour. The non-trivial experimental challenge is how to promote the use of heuristics in the laboratory. Scholars have developed four different techniques: time pressure, ego depletion, cognitive load, and conceptual primes of intuition². I review them below:

- When people have little time to think about the details of a decision problem, they might be more likely to rely on general heuristics. Therefore, putting people under *time pressure* might increase their reliance on heuristics.
- When people are depleted of their self-control, they might lose their ability to calculate the details of the decision problem at hand and, therefore, become more likely to use general heuristics. Self-control can be depleted through an *ego depletion* task, such as the Stroop task or the *e-hunting* task, or, in general, through any task that requires the use of self-control.
- When people's working memory is reduced by a concurrent task, their ability to make the complex reasoning needed to evaluate the situation they are facing might be reduced as well, making them more likely to follow simple heuristics. Working

² To be precise, there is also a fifth technique: neurostimulation. Neurostimulation methods come from the idea that high-level, reflective reasoning comes primarily from a specific brain area, the right dorsolateral prefrontal cortex (rDLPFC). Therefore, deactivating this area, using transcranial magnetic stimulation or transcranial direct current stimulation, might make people more likely to follow their heuristics. However, in this article, I decided not to focus on this method because, to the best of my knowledge, there are no studies testing the effect of neurostimulation of the rDLPFC on cooperative behaviour using prisoner's dilemmas or public goods games. There is only one study, but it uses an asymmetric public goods game (Li et al. 2018). I hope that future work can fill this gap.

memory can be depleted using *cognitive load* tasks, such as keeping in mind a long sequence of numbers (typically seven).

- *Conceptual primes of intuition* refer to a class of nudges that promote reliance on intuitive thinking. These nudges can be implicit or explicit. An instance of an implicit nudge could be that people, before playing a social dilemma, are given a set of letters that they can use to form words, and these words are related to intuitive decision making, (e.g., “intuition”, “emotion”, or “quick”). An explicit nudge could be to ask people to follow their intuition or their emotion, or to write about a time of their life in which following their intuitions worked out well.

It is important to note that none of these methods is perfect. Time pressure has been criticised because it is usually too long to eliminate reflective reasoning and access quick heuristics (Libet 2009; Soon et al. 2008). Ego depletion may even do something fundamentally different from deactivating reflective reasoning and activating reflexive reactions; moreover, the very basic assumption that self-control draws on a limited resource also came under scrutiny (Inzlicht et al. 2014). Cognitive load tasks might interact with the primary task, while conceptual primes, especially the explicit ones, may generate experimenter demand effect (Rand 2016).³

Being aware of the limitations of these experimental manipulations, scholars have turned to meta-analytic techniques to find out whether, overall (i.e., putting all the studies together, regardless of the cognitive manipulation being used), intuition favours cooperative behaviour. An earlier meta-analysis found a positive effect of promoting intuition on cooperation (Rand 2016). However, this result was later criticised by another meta-analysis, which found a null effect (Kvarven et al. 2019), which in turn was criticised by a third meta-analysis, which replicated the original positive effect (Rand 2019). The debate about whether promoting intuition increases cooperative behaviour is still ongoing (see Capraro (2019) for a review). However, there is a result that has been consistently found in all meta-analyses: explicit primes of emotions increase cooperative behaviour. To make an example, an explicit message (shown to participants before making their decision) used to prime reliance on emotions is the following:

Sometimes people make decisions by using feeling and relying on their emotion. Other times, people make decisions by using logic and relying on their reason.

Many people believe that emotion leads to good decision-making. When we use feelings, rather than logic, we make emotionally satisfying decisions.

Please make your transfer decision by relying on emotion, rather than reason.

This prime was initially introduced by Levine et al. (2018) and shown to increase cooperative behaviour in the prisoner’s dilemma. More recently, it has been applied also to other contexts. For example, it has been shown to reduce speciesism, that one can interpret as a form of cooperation between humans and non-human animals (Caviola and Capraro 2020). However, this very same prime has also been shown to *reduce* intentions

³ For completeness, I mention that also neurostimulation tools have been criticised, as they are usually applied over the brain area of interest. This implicitly assumes that the stimulus spreads uniformly towards the target area. However, this is generally not true, but depends on the topography of the cortical surface, which, in some cases, can even reverse the polarity of the stimulus (Berker et al., 2013; Rahman et al., 2013).

to wear a face mask during the COVID-19 pandemic (Barcelo & Capraro, 2021). This raises an issue that I think it deserves attention. While previous research has focused on “general cooperation” using stylised games, it is possible that particular forms of cooperation, especially those we are unfamiliar with, may require reflective reasoning.

Another important point to reflect upon is that the aforementioned work regards the effect of “general emotions” on cooperation. Specific emotions may affect cooperation in different ways, depending on the emotions themselves. For example, Polman and Kim (2013) found that inducing anger decreases cooperation in the public goods game, while inducing disgust increases cooperation. Motro et al. (2017) found that inducing anger decreases cooperation in a prisoner’s dilemma, but only when the other participant was angry as well. Chierchia et al. (2021) found that inducing fear increased cooperation compared to inducing anger, but none of them was different from the control condition.

4.2 Moral Preferences

Another account that has been proposed is the morality preferences hypothesis, according to which cooperative behaviour in one-shot and anonymous interactions is primarily driven by moral preferences for doing the right thing (Capraro and Perc 2021). This account is not mutually exclusive with the social heuristics hypothesis because personal norms – internal standards about what is right or wrong – may come from the internalisation of behaviours that have been learned in everyday interactions.

The experimental evidence on the morality preferences hypothesis began with a paper by Bizziou van Pol et al. (2015), which found that cooperative behaviour in the one-shot prisoner’s dilemma correlates positively with honest behaviour in a sender-receiver game where a sender can send a dishonest message and increase both his payoff and that of the receiver. Therefore, the honest choice in this game is considered as a measure of moral preferences beyond monetary outcomes. The fact that this choice correlates with cooperative behaviour suggests that cooperative behaviour is partly driven by moral preferences beyond monetary outcomes. A subsequent article by Kimbrough and Vostroknutov (2016) showed that cooperative behaviour in the public goods game correlates with norm-following in a task where subjects, moving a virtual person on a computer screen using a keyboard, get to decide how long to stand at a virtual traffic-light which displays the red colour, losing time and money, in a context where they receive no negative payoff if they just cross the road disregarding the traffic-light. The decision in this norm-following task is anonymous, has nothing to do with social interactions, and does not have any material consequences. Therefore, although Kimbrough and Vostroknutov (2016) do not mention personal norms in their paper, I think it is fair to consider their task as a measure of people’s propensity to follow their personal norms. In this light, their results may be interpreted as providing further support for the assumption that cooperative behaviour in one-shot and anonymous social dilemmas is partly driven by moral preferences. Then, Capraro and Rand (2018) demonstrated that cooperative behaviour in the prisoner’s dilemma correlates with the moral choice in the trade-off game, regardless of the trade-off game frame, thus providing additional evidence that

cooperative behaviour is partly driven by moral preferences beyond monetary outcomes.⁴ Recently, Bašić and Verrina (2021) provided evidence that personal norms are complementary to social norms in predictive cooperative behaviour, and Catola et al. (2021) showed that both social and personal norms are correlated with cooperative behaviour in the public goods game, but the predictive power of personal norms is higher.

This line of research suggests that nudges that make the personal norm salient might be effective at increasing cooperative behaviour. However, to the best of my knowledge, only two papers explored this question.⁵

Capraro et al. (2019) explored the effect of two norm-nudges, one based on the injunctive norm, and one based on the personal norm. Specifically, before playing a prisoner's dilemma, participants answered one of the following questions:

Personal norm question: "*What do you personally think is the morally right thing to do in this situation?*".

Injunctive norm question: "*What do you think your society considers to be the morally right thing to do in this situation?*".

After answering this question, participants played a one-shot prisoner's dilemma. The results showed that both messages increased cooperative behaviour compared to a baseline, where participants made their decision without being asked any question. Moreover, the effect of the two nudges was similar.

Mieth et al. (2021) tested the effect of moral labels on cooperative behaviour in the prisoner's dilemma. They found that labelling the available options with morally loaded words, such as "I cooperate" vs. "I cheat", increases cooperation, compared to using neutral labels "Option A" vs. "Option B".

Thus, these two studies provide evidence that making the personal norm salient tend to increase cooperative behaviour. However, it seems that also making the injunctive norm salient increases cooperative behaviour to a similar extent, suggesting that also injunctive norms may play an important role in determining one-shot and anonymous cooperation. It might be possible that different people are nudged by different norms; for example, it could be that people high in internalised moral identity tend to react to personal norm-nudges, while people high in symbolised moral identity tend to react to injunctive norm-nudges.⁶ In general, studying the role of potential moderators could be a promising route for future research. Moreover, perhaps surprisingly, I could not find any work testing the effective of making the descriptive norm salient.⁷ One can imagine that this would also increase cooperative behaviour via conditional cooperation, but it is

⁴ I refer to Capraro, Halpern and Perc (in press) for a review article providing many examples of situations in which people's behaviour cannot be explained using outcome-based utility functions but require language-based utility functions; moral preferences can be seen as particular language-based utility functions, where morally loaded language carries the moral utility of an action.

⁵ There is also one work exploring the effect of moral messages in the iterated prisoner's dilemma. Dal Bó and Dal Bó (2014) found that making participants read the Golden Rule increases cooperation in the iterated prisoner's dilemma, although the effect vanishes after a few rounds.

⁶ Internalised moral identity measures the extent to which being moral is important to one's self-concept, while symbolised moral identity measures the extent to which people care about looking moral (Aquino and Reed 2002).

⁷ The descriptive norm represents what other people actually do (Cialdini et al. 1990).

an open question if this is actually the case and, if so, how the magnitude of this effect compares with the magnitude of the effects of the other norm-nudges.

5 Conclusion

In this article, I reviewed the main mechanisms and interventions that are known to promote cooperative behaviour in social dilemma games. I summarise them in Table 1. I also highlighted some open questions that I hope can be answered in future work. I summarise them in Table 2.

Table 1. Summary of the mechanisms and interventions that are known to increase cooperative behaviour in social dilemmas.

Iterated or non-anonymous games
<ul style="list-style-type: none"> • <i>Genetic relatedness</i> increases cooperative behaviour • <i>Future interactions</i> increase cooperative behaviour • <i>Public reputation</i> increases cooperative behaviour • <i>Networked interactions</i> increase cooperative behaviour, especially if one can break old links and create new ones • <i>Group competition</i> increases cooperative behaviour withing groups • <i>Reducing the cost</i> of cooperation or <i>increasing its benefit</i> increases cooperative behaviour • <i>Larger groups</i> are <i>more cooperative</i> in the <i>public goods game</i>, but <i>less cooperative</i> in the <i>prisoner's dilemma</i> • <i>Punishment</i> or <i>reward</i> increases cooperative behaviour
One-shot and anonymous games
<ul style="list-style-type: none"> • <i>Reducing the cost</i> of cooperation or <i>increasing its benefit</i> increases cooperative behaviour • <i>Larger groups</i> are <i>more cooperative</i> in the <i>public goods game</i>, but <i>less cooperative</i> in the <i>prisoner's dilemma</i> • <i>Cues</i> suggesting that <i>interactions might be repeated</i> or <i>non-anonymous</i> increase cooperative behaviour • <i>Inducing general emotions</i> increases cooperative behaviour • <i>Personal norm-nudges</i> increase cooperative behaviour • <i>Injunctive norm-nudges</i> increase cooperative behaviour

Table 2. Open questions.

Open questions
<ul style="list-style-type: none"> • What is the effect of group size on cooperation in the piecewise linear-then-constant public goods game? If inverted-U, is it possible to estimate the size of the group that produces the largest public good? • Does intuition promote cooperation? • Do specific forms of cooperation – arguably those we are not familiar with – require deliberative thinking? • Which specific emotions promote cooperation? Which undermine it? • What is the effect of neurostimulation of the dorsolateral prefrontal cortex on cooperative behaviour? • Do descriptive norm nudges increase cooperative behaviour? How does their effect compare with the effect of personal and injunctive norm-nudges? • Are people who respond to personal norm nudges somehow different from people who respond to injunctive norm nudges?

References

- Aquino, K., Reed, A., II.: The self-importance of moral identity. *J. Pers. Soc. Psychol.* **83**, 1423–1440 (2002)
- Barcelo, H., Capraro, V.: Group size effect on cooperation in one-shot social dilemmas. *Sci. Rep.* **5**, 7937 (2015)
- Bašić, Z., Verrina, E.: Personal norms—and not only social norms—shape economic behavior. MPI Collective Goods Discussion Paper, (2020/25)
- Basu, K.: The traveler’s dilemma: paradoxes of rationality in game theory. *Am. Econ. Rev.* **84**, 391–395 (1994)
- de Berker, A.O., Bikson, M., Bestmann, S.: Predicting the behavioral impact of transcranial direct current stimulation: Issues and limitations. *Front. Hum. Neurosci.* **7**, 613 (2013)
- Bertrand, J.: Book review of *theorie mathematique de la richesse social* and of *recherches sur les principes mathematiques de la theorie des richesses*. *J. des Savants* (1883)
- Bilancini, E., Boncinelli, L.: The co-evolution of cooperation and defection under local interaction and endogenous network formation. *J. Econ. Behav. Organ.* **70**, 186–195 (2009)
- Bilancini, E., Boncinelli, L., Wu, J.: The interplay of cultural intolerance and action-assortativity for the emergence of cooperation and homophily. *Eur. Econ. Rev.* **102**, 1–18 (2018)
- Binmore, K.: Modeling rational players: part I. *Econ. Philos.* **3**, 179–214 (1987)
- Biziou van Pol, L., Haenen, J., Novaro, A., Occhipinti-Liberman, A., Capraro, V.: Does telling white lies signal pro-social preferences? *Judgm. Decis. Mak.* **10**, 538–548 (2015)
- Böhm, R., Rockenbach, B.: The inter-group comparison – intra-group cooperation hypothesis: comparisons between groups increase efficiency in public goods provision. *PLoS ONE* **8**, e56152 (2013)
- Bolton, G.E., Katok, E., Ockenfels, A.: Cooperation among strangers with limited information about reputation. *J. Public Econ.* **89**, 1457–1468 (2005)
- Boyd, R., Gintis, H., Bowles, S., Richerson, P.J.: The evolution of altruistic punishment. *Proc. Natl. Acad. Sci.* **100**, 3531–3535 (2003)
- Capraro, V.: The dual-process approach to human sociality: a review. Available at SSRN 3409146 (2019)
- Capraro, V., Barcelo, H.: Group size effect on cooperation in one-shot social dilemmas II: curvilinear effect. *PLoS ONE* **10**, e0131419 (2015)

- Capraro, V., Barcelo, H.: Punishing defectors and rewarding cooperators: do people discriminate between genders? *J. Econ. Sci. Assoc.* **7**(1), 19–32 (2021a). <https://doi.org/10.1007/s40881-021-00099-4>
- Capraro, V., Barcelo, H.: Telling people to “rely on their reasoning” increases intentions to wear a face covering to slow down COVID-19 transmission. *Appl. Cogn. Psychol.* **35**, 693–699 (2021b)
- Capraro, V., Cococcioni, G.: Rethinking spontaneous giving: extreme time pressure and ego-depletion favor self-regarding reactions. *Sci. Rep.* **6**, 27219 (2016)
- Capraro, V., Giardini, F., Vilone, D., Paolucci, M.: Partner selection supported by opaque reputation promotes cooperative behavior. *Judgm. Decis. Mak.* **11**, 589–600 (2016)
- Capraro, V., Halpern, J. Y., Perc, M.: From outcome-based to language-based preferences. *J. Econ. Literature* (in press)
- Capraro, V., Jagfeld, G., Klein, R., Mul, M., de Pol, I.V.: Increasing altruistic and cooperative behaviour with simple moral nudges. *Sci. Rep.* **9**, 11880 (2019)
- Capraro, V., Jordan, J.J., Rand, D.G.: Heuristics guide the implementation of social preferences in one-shot Prisoner’s Dilemma experiments. *Sci. Rep.* **4**, 6790 (2014)
- Capraro, V., Perc, M.: Mathematical foundations of moral preferences. *J. R. Soc. Interface* **18**, 20200880 (2021)
- Capraro, V., Rand, D.G.: Do the right thing: experimental evidence that preferences for moral behavior, rather than equity and efficiency per se, drive human prosociality. *Judgm. Decis. Mak.* **13**, 99–111 (2018)
- Carroll, J.W.: Iterated N-player prisoner’s dilemma games. *Philos. Stud.: Int. J. Philos. Anal. Tradit.* **53**, 411–415 (1988)
- Catola, M., D’Alessandro, S., Guarnieri, P., Pizziol, V.: Personal norms in the online public good game. *Econ. Lett.* **207**, 110024 (2021)
- Caviola, L., Capraro, V.: Liking but devaluing animals: emotional and deliberative paths to speciesism. *Soc. Psychol. Pers. Sci.* **11**, 1080–1088 (2020)
- Chierchia, G., Parianen Lesemann, F.H., Snower, D., Singer, T.: Cooperation across multiple game theoretical paradigms is increased by fear more than anger in selfish individuals. *Sci. Rep.* **11**, 9351 (2021)
- Cialdini, R.B., Reno, R.R., Kallgren, C.A.: A focus theory of normative conduct: recycling the concept of norms to reduce littering in public places. *J. Pers. Soc. Psychol.* **58**, 1015–1026 (1990)
- Dal Bó, P., Fréchet, G.R.: The evolution of cooperation in infinitely repeated games: experimental evidence. *Am. Econ. Rev.* **101**, 411–429 (2011)
- Dal Bó, E., Dal Bó, P.: Do the right thing: the effects of moral suasion on cooperation. *J. Public Econ.* **117**, 28–38 (2014)
- Duffy, J., Ochs, J.: Cooperative behavior and the frequency of social interaction. *Games Econ. Behav.* **66**, 785–812 (2009)
- Dunbar, R.I., Marriott, A., Duncan, N.D.: Human conversational behavior. *Hum. Nat.* **8**, 231–246 (1997)
- Engel, C., Zhurakhovska, L.: When is the risk of cooperation worth taking? The prisoner’s dilemma as a game of multiple motives. *Appl. Econ. Lett.* **23**, 1157–1161 (2016)
- Erev, I., Bornstein, G., Galili, R.: Constructive intergroup competition as a solution to the free rider problem: a field experiment. *J. Exp. Soc. Psychol.* **29**, 463–478 (1993)
- Evans, J.S.B.T., Stanovich, K.E.: Dual-process theories of higher cognition: advancing the debate. *Perspect. Psychol. Sci.* **8**, 223–241 (2013)
- Fehl, K., van der Post, D.J., Semman, D.: Co-evolution of behaviour and social network structure promotes human cooperation. *Ecol. Lett.* **14**, 546–551 (2011)
- Fehr, E., Fischbacher, U.: The nature of human altruism. *Nature* **425**, 785–791 (2003)

- Fehr, E., Gächter, S.: Cooperation and punishment in public goods experiments. *Am. Econ. Rev.* **90**, 980–994 (2000)
- Fehr, E., Gächter, S.: Altruistic punishment in humans. *Nature* **415**, 137–140 (2002)
- Fischbacher, U., Gächter, S., Fehr, E.: Are people conditionally cooperative? Evidence from a public goods experiment. *Econ. Lett.* **71**, 397–404 (2001)
- Fudenberg, D., Rand, D.G., Dreber, A.: Slow to anger and fast to forgive: cooperation in an uncertain world. *Am. Econ. Rev.* **102**, 720–749 (2012)
- Gracia-Lázaro, C., et al.: Heterogeneous networks do not promote cooperation when humans play a Prisoner's Dilemma. *Proc. Natl. Acad. Sci.* **109**, 12922–12926 (2012)
- Grujić, J., Eke, B., Cabrales, A., Cuesta, J.A., Sánchez, A.: Three is a crowd in iterated prisoner's dilemmas: experimental evidence on reciprocal behavior. *Sci. Rep.* **2**, 638 (2012b)
- Grujić, J., Fosco, C., Araujo, L., Cuesta, J.A., Sánchez, A.: Social experiments in the mesoscale: humans playing a spatial prisoner's dilemma. *PLoS ONE* **5**, e13749 (2010)
- Grujić, J., Röhl, T., Semmann, D., Milinski, M., Traulsen, A.: Consistent strategy updating in spatial and non-spatial behavioral experiments does not promote cooperation in social networks. *PLoS ONE* **7**, e47718 (2012a)
- Gunthorsdottir, A., Houser, D., McCabe, K.: Disposition, history and contributions in public goods experiments. *J. Econ. Behav. Organ.* **62**, 304–315 (2007)
- Gunthorsdottir, A., Rapoport, A.: Embedding social dilemmas in intergroup competition reduces free-riding. *Organ. Behav. Hum. Decis. Process.* **101**, 184–199 (2006)
- Hamburger, H.: N-person prisoner's dilemma. *J. Math. Sociol.* **3**, 27–48 (1973)
- Hamilton, W.D.: The genetical evolution of social behaviour. *J. Theor. Biol.* **7**, 17–52 (1964)
- Heckathorn, D.D.: The dynamics and dilemmas of collective action. *Am. Soc. Rev.* **61**, 250–277 (1996)
- Herrmann, B., Thoni, C., Gächter, S.: Antisocial punishment across societies. *Science* **319**, 1362–1367 (2008)
- Inzlicht, M., Schmeichel, B.J., Macrae, C.N.: Why self-control seems (but may not be) limited. *Trends Cogn. Sci.* **18**, 127–133 (2014)
- Isaac, R.M., Walker, J.M., Williams, A.W.: Group size and the voluntary provision of public goods: experimental evidence utilizing large groups. *J. Public Econ.* **54**(1), 1–36 (1994)
- Kahneman, D.: *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux (2011)
- Kimbrough, E.O., Vostroknutov, A.: Norms make preferences social. *J. Eur. Econ. Assoc.* **14**, 608–638 (2016)
- Kocher, M.G., Cherry, T., Kroll, S., Netzer, R.J., Sutter, M.: Conditional cooperation on three continents. *Econ. Lett.* **101**, 175–178 (2008)
- Kvarven, A., et al.: The intuitive cooperation hypothesis revisited: a meta-analytic examination of effect-size and between-study heterogeneity. *J. Econ. Sci. Assoc.* **6**, 26–42 (2019)
- Levine, E., Barasch, A., Rand, D.G., Berman, J.Z., Small, D.A.: Signaling emotion and reason in cooperation. *J. Exp. Psychol. Gen.* **5**, 702–719 (2018)
- Li, J., Liu, X., Yin, X., Wang, G., Niu, X., Zhu, C.: Transcranial direct current stimulation altered voluntary cooperative norms compliance under equal decision-making power. *Front. Hum. Neurosci.* **12**, 265 (2018)
- Libet, B.: *Mind Time: The Temporal Factor in Consciousness*. Harvard University Press, Cambridge (2009)
- Madsen, E.A., et al.: Kinship and altruism: a cross-cultural experimental study. *Br. J. Psychol.* **98**, 339–359 (2007)
- Mieth, L., Buchner, A., Bell, R.: Moral labels increase cooperation and costly punishment in a Prisoner's Dilemma game with punishment option. *Sci. Rep.* **11**, 10221 (2021)
- Milinski, M., Semmann, D., Krambeck, H.J.: Reputation helps solve the 'tragedy of the commons.' *Nature* **415**, 424–426 (2002)

- Motro, D., Kugler, T., Connolly, T.: Back to the basics: how feelings of anger affect cooperation. *Int. J. Conf. Manag.* **27**, 523–546 (2016)
- Murnighan, J.K., Roth, A.E.: Expecting continued play in prisoner's dilemma games: a test of several models. *J. Conflict Resolut.* **27**, 279–300 (1983)
- Nowak, M.A.: Five rules for the evolution of cooperation. *Science* **314**, 1560–1563 (2006)
- Nowak, M.A., Sigmund, K.: Evolution of indirect reciprocity by image scoring. *Nature* **393**, 573–577 (1998)
- Ostrom, E., Walker, J., Gardner, R.: Covenants with and without a sword: self-governance is possible. *Am. Polit. Sci. Rev.* **86**, 404–417 (1992)
- Ohtsuki, H., Hauert, C., Lieberman, E., Nowak, M.A.: A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502–505 (2006)
- Perc, M., Jordan, J.J., Rand, D.G., Wang, Z., Boccaletti, S., Szolnoki, A.: Statistical physics of human cooperation. *Phys. Rep.* **687**, 1–51 (2017)
- Pereda, M., Capraro, V., Sánchez, A.: Group size effects and critical mass in public goods games. *Sci. Rep.* **9**, 5503 (2019)
- Peysakhovich, A., Nowak, M.A., Rand, D.G.: Humans display a “cooperative phenotype” that is domain general and temporally stable. *Nat. Commun.* **5**, 4939 (2014)
- Pfeiffer, T., Tran, L., Krumme, C., Rand, D.G.: The value of reputation. *J. R. Soc. Interface* **9**, 2791–2797 (2012)
- Polman, E., Kim, S.H.: Effects of anger, disgust, and sadness on sharing with others. *Pers. Soc. Psychol. Bull.* **39**, 1683–1692 (2013)
- Puurttinen, M., Mappes, T.: Between-group competition and human cooperation. *Proc. Royal Soc. B: Biol. Sci.* **276**, 355–360 (2009)
- Rahman, A., Reato, D., Arlotti, M., Gasca, F., Datta, A., Parra, L.C., et al.: Cellular effects of acute direct stimulation: somatic and synaptic terminal effects. *J. Physiol.* **591**, 2563–2578 (2013)
- Rand, D.G., Arbesman, S., Christakis, N.A.: Dynamic social networks promote cooperation in experiments with humans. *Proc. Natl. Acad. Sci.* **108**, 19193–19198 (2011)
- Rand, D.G., Dreber, A., Ellingsen, T., Fudenberg, D., Nowak, M.A.: Positive interactions promote public cooperation. *Science* **325**, 1272–1275 (2009)
- Rand, D.G., Nowak, M.A., Fowler, J.H., Christakis, N.A.: Static network structure can stabilize human cooperation. *Proc. Natl. Acad. Sci.* **111**, 17093–17098 (2014a)
- Rand, D.G.: Cooperation, fast and slow: Meta-analytic evidence for a theory of social heuristics and self-interested deliberation. *Psychol. Sci.* **27**, 1192–1206 (2016)
- Rand, D.G.: Intuition, deliberation, and cooperation: Further meta-analytic evidence from 91 experiments on pure cooperation (2019). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3390018
- Rand, D.G., Nowak, M.A.: Human cooperation. *Trends Cogn. Sci.* **17**, 413–425 (2013)
- Rand, D.G., et al.: Social heuristics shape intuitive cooperation. *Nat. Commun.* **5**, 3677 (2014b)
- Rapoport, A., Chammah, A.M.: Prisoner's Dilemma: A Study in Conflict and Cooperation (Vol. 165). University of Michigan press, Ann Arbor (1965)
- Richerson, P., et al.: Cultural group selection plays an essential role in explaining human cooperation: a sketch of the evidence. *Behav. Brain Sci.* **39**, 1–68 (2016)
- Rockenbach, B., Milinski, M.: The efficient interaction of indirect reciprocity and costly punishment. *Nature* **444**, 718–723 (2006)
- Roth, A.E., Murnighan, J.K.: Equilibrium behavior and repeated play of the prisoner's dilemma. *J. Math. Psychol.* **17**, 189–198 (1978)
- Seinen, I., Schram, A.: Social status and group norms: indirect reciprocity in a repeated helping experiment. *Eur. Econ. Rev.* **50**, 581–602 (2006)
- Sommerfeld, R.D., Krambeck, H.J., Semmann, D., Milinski, M.: Gossip as an alternative for direct observation in games of indirect reciprocity. *Proc. Natl. Acad. Sci.* **104**, 17435–17440 (2007)

- Soon, C.S., Brass, M., Heinze, H.J., Haynes, J.D.: Unconscious determinants of free decisions in the human brain. *Nat. Neurosci.* **11**, 543–545 (2008)
- Suri, S., Watts, D.J.: Cooperation and contagion in web-based, networked public goods experiments. *PLoS ONE* **6**, e16836 (2011)
- Tan, J.H., Bolle, F.: Team competition and the public goods game. *Econ. Lett.* **96**, 133–139 (2007)
- Tomasello, M., Carpenter, M., Call, J., Behne, T., Moll, H.: Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* **28**, 675–691 (2005)
- Traulsen, A., Nowak, M.A.: Evolution of cooperation by multilevel selection. *Proc. Natl. Acad. Sci.* **103**, 10952–10955 (2006)
- Traulsen, A., Semmann, D., Sommerfeld, R.D., Krambeck, H.J., Milinski, M.: Human strategy updating in evolutionary games. *Proc. Natl. Acad. Sci.* **107**, 2962–2966 (2010)
- Wang, J., Suri, S., Watts, D.J.: Cooperation and assortativity with dynamic partner updating. *Proc. Natl. Acad. Sci.* **109**, 14363–14368 (2012)
- Yamagishi, T.: The provision of a sanctioning system as a public good. *J. Pers. Soc. Psychol.* **51**, 110–116 (1986)
- Yang, W., et al.: Nonlinear effects of group size on collective action and resource outcomes. *Proc. Natl. Acad. Sci.* **110**, 10916–10921 (2013)
- Yao, X., Darwen, P.J.: An experimental study of N-person iterated prisoner’s dilemma games. *Informatica* **18**, 435–450 (1994)