

**Slowing COVID-19 transmission as a social dilemma: Lessons  
for government officials from interdisciplinary research on  
cooperation**

Tim Johnson\*  
Willamette University

Christopher T. Dawes  
New York University

James H. Fowler  
University of California, San Diego

Oleg Smirnov  
Stony Brook University

\*Address correspondence to Tim Johnson at [tjohnson@willamette.edu](mailto:tjohnson@willamette.edu).

**Transparency:** This article solely reviews past literature, thus it contains no data, materials, or preregistration to make publicly available.

---

## Slowing COVID-19 transmission as a social dilemma: Lessons for government officials from interdisciplinary research on cooperation

### **Abstract:**

To reduce transmission of COVID-19, public officials must help their communities resolve a series of novel social dilemmas. For instance, when social distancing becomes widespread, the likelihood of COVID-19 exposure decreases, thus tempting individuals to leave their homes while others stay sheltered. Yet, if all indulge that temptation, then rates of transmission will increase: everyone would have fared better by cooperatively staying at home. Past research has studied such social dilemmas to understand why cooperation occurs despite incentives that conspire against it. In this narrative review, we select relevant insights from this literature to inform COVID-19 response and we structure those insights around the response stages that government officials face. Together, the measures that we identify can ameliorate the social dilemmas born from the COVID-19 pandemic.

**Keywords:** Social dilemma, Cooperation, COVID-19, Free-riding

To reduce transmission of COVID-19, public officials must help their communities resolve a series of novel social dilemmas. For instance, COVID-19 increases the risk of mortality by very little for any healthy individual. However, social distancing and staying at home impose personal costs. The benefits of those precautionary measures accrue to more-vulnerable individuals and the community writ large by lessening the chance that a wider outbreak overwhelms hospitals (Ruoran et al., 2020)—a major concern if transmission does not decrease. Given these personal costs and diffuse benefits, individuals face the lure of living their lives as normal, while others pay the costs of staying at home or social distancing. Anecdotal evidence of crowded bars (Rogers & Marshall, 2020), jet-setters enjoying travel at low cost (Hoffower, 2020), and young people mingling in defiance of their elders' best interests (Pancevski & Meichtry, 2020) suggests that the above stylized model comports with reality.

Models of other dilemmas resulting from the COVID-19 crisis exhibit similar features. For example, as social distancing becomes widespread, the likelihood of COVID-19 exposure decreases, thus tempting individuals to leave their homes while others stay sheltered. Yet, if all indulge that temptation, then rates of transmission increase: everyone would have fared better by cooperatively staying at home. Likewise, consider a challenge facing government officials. Government officials can facilitate a more-effective response if they cooperatively distribute medical supplies across jurisdictions, but they and their home jurisdictions become better situated by hoarding supplies while others share resources. If all adopt that self-interested tactic, however, a coherent response fails, prolonging transmission and leading all to wish they had incurred the less-significant costs of a cooperative response. In each of these stylized models, personal incentives compel individuals to act against group

welfare, despite an ultimate fate that leaves them wishing they had acted in the community's interest.

Each model, in other words, echoes the incentive structure of a social dilemma (Dawes, 1980; Nowak & Highfield, 2011). In a social dilemma, individuals can produce some benefit that all will enjoy, but they must absorb a personal cost to do so. Individuals therefore maximize their personal welfare via free riding, which earns them the benefits of cooperation without tolerating its costs. Cooperators endure exploitation in such instances, thus providing reason for them to resist cooperation in the first place and add themselves to the ranks of defectors. This dynamic generates a Pareto-suboptimal state of universal defection, however, such that all would have preferred unanimous cooperation.

For roughly seventy years, a cross-disciplinary literature in anthropology, biology, economics, political science, psychology, and sociology has studied variants of the social dilemma to understand why cooperation occurs despite incentives that conspire against it (Dawes, 1980; Poundstone, 1992). This literature offers many “solutions” to the social dilemma—that is, it offers guidance on how to promote personally costly behaviors that yield benefits to the community despite the enticing prospect of free riding. In this article, we select relevant insights from this literature that offer guidance on how public officials around the globe can address the social dilemma that rests at the heart of slowing the transmission of the coronavirus. Other researchers have concurrently reviewed the wider literature on how the social and behavioral sciences can inform responses to the coronavirus (Lunn et al., 2020; Van Bavel et al., 2020); our narrative review pursues the narrower goal of focusing on how the literature on social dilemmas can contribute to the COVID-19 response. Moreover, our review not only provides guidance on the interventions that public

officials can employ to promote cooperative COVID-19 response efforts, but it also provides insight into the behaviors that public officials should expect from their communities' members when they face the social dilemmas resting at the heart of the current crisis. We hope this latter form of information helps public officials anticipate community responses and develop their own, localized solutions to the social dilemmas their jurisdictions encounter.

We organize our presentation of these insights around stages of COVID-19 response and we pair insights from the literature with the practical advice they suggest. Table 1, on the next page, summarizes our review. With this mode of organization, we hope to provide guidance for the prioritization of measures depending on the progress of transmission in a given setting. Although large urban areas already have passed early stages in which nonintrusive measures might stop transmission, remote areas may still be able to use those measures to support the cooperation needed to slow coronavirus transmission. *We emphasize that our review of the literature seeks to provide rough guidance, not the detailed plans of public health experts.* People should follow those officials' plans. As will become evident here, the research we discuss provides evidence to support the behavioral postulates on which those plans rest and to emphasize ways in which those plans can be more effective.

**Table 1.**

**Key Insights for COVID-19 Response from the Literature on Social Dilemmas**

Stage of Response	Mechanism Supporting Cooperation	Practical Measures that Use the Mechanism
Immediate Guidance	Cooperative Instincts	-Advise the public to trust its intuition about social responsibility
	Spatial and Network Insulation	-Advocate social detachment from individuals who do not cooperate with social distancing and other response measures
		-Create opportunities for remote interaction of cooperative individuals engaged in distancing
Initial Implementation		-Create additional penalties for conspicuously violating distancing rules or other response measures
	Group Identity	-Communicate that the benefits of cooperation accrue to those holding a shared, community-wide identity
Sustaining Non-Mandatory Measures	Reciprocity	-Reward cooperative acts
		-Forge reciprocal exchange programs across communities hit by the pandemic at different times
Enforcing Mandatory Measures	Decentralized Costly Punishment	-Set clear standards for what constitutes cooperative behavior and acceptable costly punishment
		-Anticipate vigilantism and remind individuals of its illegality
	Centralized Costly Punishment	-Introduce opportunities for individuals to donate specifically to policing measures related to COVID-19 response
		-Create hotlines for informing authorities about violations of COVID-19 response measures

## **Supporting Cooperative Instincts when Providing Immediate Guidance**

Over a month after the first documented case of COVID-19 in the United States (Holshue et al., 2020), public figures ranging from state health officials (Givas, 2020) to the head of a major sports league (Dawson, 2020) to the president of the United States (Superville & Miller, 2020) cautioned against “panic” as a response to COVID-19. Instead, they proposed a reasoned, methodical approach. Although it remains sensible to admonish panic that takes the form of hoarding medical equipment or threatening physical violence, encouragement of an overtly deliberative response might hinder cooperative efforts to slow coronavirus transmission.

Over the past decade, research indicates that humans instinctually cooperate in social dilemma experiments. Rand et al. examined the relationship between participants’ response times and their rates of cooperation in social dilemma experiments (Rand, Greene, & Nowak, 2012). Prior work indicates that response times in experiments correlate with automatic processing and behavioral response, whereas longer response times reflect deliberative, reasoned cognition (cf. Krajbich, Bartling, Hare, & Fehr, 2015). Performing a novel experiment conducted with participants engaged in a variant of the social dilemma known as a one-shot public goods game,<sup>1</sup> Rand et al. found a negative relationship between response times and contribution rates (Rand et al., 2012). Further study of previously unexplored data from past experiments conducted by the authors’ lab showed a similar trend. With these exploratory findings in tow, the authors designed a novel experiment that randomly primed intuition versus deliberation; they found that participants placed in the intuition condition contributed more in the public goods game. Moreover, testing the possibility that this instinct is learned through successful past cooperation, Rand et al. found

in still another study that participants who had previously played the public goods game and, thus, had earlier opportunities to reflect on it, contributed less in a subsequent public goods game (Rand et al., 2012). Further research has offered formal analyses indicating the conditions in which this behavior proves adaptive (Jagau & van Veelen, 2017) and has replicated its findings in meta-analytic studies (Rand et al., 2014) and novel replications (Isler, Maule, & Starmer, 2018). Estimates from meta-analyses suggest that forcing knee-jerk reactions to social dilemmas increases cooperation by 21.5% (Rand et al., 2014, p.3).

These results question public officials who ask members of the community to reason through their response to the coronavirus crisis in lieu of acting on instinct. If asked to reason in detail, individuals will recognize that the risk to their personal mortality is small, thus impeding their instinct to cooperate. In communities that have yet to experience cases, public officials may still advise a methodical approach, thereby inadvertently encouraging individuals to surmise that, indeed, a trip to the store or a dinner out to celebrate an anniversary increases transmission and personal risk only a tiny amount. To the community, those increases aggregate to speed transmission and impose a steep cost on the group. Current estimates place the value of social distancing at approximately \$60,000 per household (Greenstone and Nigam, 2020).

If detailed advice is necessary, public officials should state it directly (e.g., “buy groceries as you normally would”) instead of advising against specific forms of panic (e.g. “don’t panic buy”); the latter statements might induce the very behavior they reject (see Wegner et al., 1987)<sup>2</sup> and trigger methodical reasoning that promotes selfishness (Rand et al., 2014). If officials wish to slow transmission of the coronavirus, they should advise individuals to trust their instincts about what actions are socially responsible. Research on



the social dilemma indicates that those instincts will be cooperative. This advice, furthermore, will hold as new social dilemmas relevant to the COVID-19 crisis emerge, such as how to respond to slowing transmission or how to deal with resource shortages. In each instance, public officials can prepare their communities for these social dilemmas by encouraging the use of individuals' cooperative instincts.

*Lessons for public officials:*

1. Advise the public to trust its intuition about social responsibility
2. Recast generic, negative advice—such as “don't panic”—into detailed recommendations of specific, beneficial behaviors

### **Reducing Behavioral Variation by Insulating Self-organizing Pools of Cooperators**

Supporting instinctive cooperation helps cooperation take root in social dilemmas, but how can communities support the persistence of that cooperation? For instance, students in a dorm might cooperatively implement social distancing only to see from their windows an adjacent dorm throwing a party with no immediate ill effects. Individuals living their lives normally with no apparent consequences might entice individuals to deviate from their current course of cooperative activity. Not only could social mixing result in increased disease transmission, but it also might spread the non-cooperative behavior that facilitates transmission.

Such dynamics figure prominently in mathematical models studying the social dilemma (Nowak, 2006a). In those models, theoreticians investigate populations of individuals who occasionally update their behavioral strategy to cooperate or defect; the rules

for updating vary, but they almost universally depict a process in which more-successful strategies in prior time periods become more common in subsequent periods (e.g., across generations). These functions primarily aim to depict evolutionary processes, but they also can capture social imitation or learning in which individuals gravitate toward more-successful strategies (Taylor & Jonker, 1978).

In the most basic of these models, researchers assume a “well-mixed” population, meaning that any given individual in the population is equally likely to engage in a social dilemma with any other given individual. In such populations, defection rapidly proliferates throughout the population as individuals observe defectors doing better by exploiting cooperators (Nowak, 2006a).

However, models that impose a spatial or network structure that eliminates thorough mixing of the population show that cooperation can emerge even without complex behavioral strategies or institutions (Nowak, 2006b). These models assume that the probability of interaction between members of the population occur in proportion to those agents’ spatial proximity or network connections. They also hold that updating strategies results from local observation: neighbors compare their well-being against that of their neighbors and switch strategies to “keep up with the Joneses,” or they examine the payoffs enjoyed by their network connections and switch strategies only if their payoffs appear less attractive than those of their network connections. Even these minimal modeling changes show that cooperation can persist so long as prosocial members of the population can achieve a degree of isolation from defectors (Nowak, 2006b; Ohtsuki, Hauert, Lieberman, & Nowak, 2006; Santos & Pacheco, 2005; Szabó & Fath, 2007). Those cooperators come to reap the gains of cooperation, while defectors experience the suboptimal state of mutual

defection, thus making the strategy of cooperation more attractive. Although the degree of cooperators' persistence and growth due to spatial or network insulation depends on a model's parameter settings, under reasonable values the insulation of cooperators can expand the pool of cooperators from a small percent of the population to the most-common type in the population—thus implying theoretical effects sizes in excess of 100% (Nowak, 2006, pp.146-161).

These models indicate measures to support the cooperation needed to slow COVID-19 transmission. First, to insulate cooperators from defectors, a simple strategy of “walking away” from non-cooperative activities works (Aktipis, 2004); thus, officials should produce public service announcements that encourage people to detach socially from those violating social-distancing measures and engage remotely with those cooperating with social-distancing efforts. Such guidance would help individuals dispassionately cope with disagreements among close affiliates concerning distancing requirements (see, e.g., the examples in Melamed, 2020). Furthermore, the emergence of videoconferencing happy hours shows that this practice remains viable (Zetlin, 2020).

Public officials also should not shy away from policy measures to prevent cooperative households from observing defectors and the short-term pleasures they are apt to be enjoying by living their lives as normal. For example, mayors in spring-break destinations who observed an influx of visitors in recent weeks and pronounced that those tourists should go home (see Acker, 2020) acted in a manner that likely insulated cooperators in their communities. That is, in addition to protecting their communities from increased virus transmission, these mayors sheltered their communities from behavioral transmission: viewing individuals who merrily engaged in standard vacation activities may have led

cooperative households to deviate from their social austerity. The theoretical literature on social dilemmas indicates that these sequestering efforts can help pools of cooperators persist (Nowak, 2006a, 2006b). Indeed, officials ought to consider additional penalties for actions that violate official distancing measures in highly visible ways that might encourage non-cooperation by others.

Such measures, moreover, might be particularly important for promoting cooperation among young teenagers. Gutiérrez-Roig, Gracia-Lázaro, Perelló, Moreno, & Sánchez (2014, SI, p.3, fig.3) found that prior behavior poorly predicts the cooperation of young teenagers (ages 10-16 years old), despite forecasting adult's cooperation well. Instead, young teenagers' likelihood of cooperation grows with the number of cooperative individuals to which they are exposed (Gutiérrez-Roig et al., 2014, p.4, fig.2). As a result, insulating cooperators from observing individuals who defy COVID-19 response measures might have its greatest effect on these socially impressionable young teenagers. Other age groups—including the young adult “spring breakers” who received much attention in the U.S.—exhibit more-stable, common patterns of cooperation that change on the margins due to alternative mechanisms for promoting cooperation, such as group identity and reciprocity (Gutiérrez-Roig et al., 2014, p.6).

*Lessons for public officials:*

1. Advocate social detachment from individuals who do not cooperate with social distancing and other response measures
2. Create opportunities for remote interaction among cooperative individuals engaged in distancing

3. Create additional penalties for *conspicuously* violating distancing rules or other response measures

### **Expanding the Pool of Cooperators through Appeals to Group Identity**

In the face of a growing number of infected individuals in his city, New York City Mayor Bill de Blasio advised that “[t]hings will continue to get worse before they get better; but New Yorkers are strong and I know we can get through this together” (Press Office of New York City Hall, 2020). By offering a brief allusion to New Yorkers’ common affiliation, de Blasio illustrated a tactic for instilling cooperation that has received attention in the literature on social dilemmas—appeals to group identity.

Early experimental research on the social dilemma found that groups of subjects that engaged in conversation often would make promises to cooperate, but these promises only appeared to influence choices in the social dilemma when they were made universally (Dawes, Van De Kragt, & Orbell, 1988). Researchers determined, after refuting alternative hypotheses, that the promises created a sense of group identity that promoted cooperation (Dawes et al., 1988). Later studies have found similar evidence of pro-sociality among in-group members (Bernhard, Fischbacher, & Fehr, 2006), as anticipated by mathematical models showing that the coupling of in-group cooperation and out-group aggression proves robust in evolutionary competition (Choi & Bowles, 2007). Subsequent empirical studies have found mixed evidence for the effect of group identity on unconditional in-group cooperation (Yamagishi & Mifune, 2016), but the bulk of the evidence appears to establish that group identity does not deteriorate in-group cooperation. In the instances in which it does work, group identity has a marked effect on cooperation: Dawes et al. (1988, p.91)

found that in experimental conditions in which group identity could influence cooperation, rates of cooperation increased from an average of 31% to a mean of 69%—a 38 percentage point jump that translates into a percent increase of 122.6%.

Accordingly, efforts to appeal to group identity may help in widening the pool of individuals cooperating with measures to slow the spread of COVID-19. By appealing to a common group identity, public officials might encourage defectors who share that common group identity to alter their behavior and become more cooperative.

*Lessons for public officials:*

1. Communicate that the benefits of cooperation accrue to those holding a shared, community-wide identity
2. Oaths or pledges can bolster affiliation with a community-wide identity that enhances cooperation (as in Dawes et al., 1988)

### **Promoting decentralized, non-mandatory cooperation via reciprocity**

Initial efforts to slow transmission of COVID-19 have focused on non-mandatory recommendations to engage in social distancing and self-quarantine. Without legal force behind them, these measures require voluntary cooperation and mild, informal policing of defection. Since the origins of research on the social dilemma, scholars have studied mechanisms of that variety.

Indeed, in the first experimental study of the social dilemma<sup>3</sup> by Merrill Flood and Melvin Dresher (De Herdt, 2003), one of the experiment's two participants sought to "train" the other participant on how to cooperate. As mathematicians, Flood's and Dresher's lack of

training in laboratory experiments led them to invite two of their colleagues—the economist Armen Alchian and their Rand Institute colleague John Williams—to play the social dilemma 100 times while keeping a journal of their thoughts during game play (De Herdt, 2003). Williams' journal indicates that he sought to guide Alchian toward cooperation via a type of reciprocity—start out with cooperation, tolerate some free-riding, but meet sustained defection with defection.

Theoretical research on altruism and the social dilemma later showed that reciprocity can sustain the evolution of cooperation (Axelrod, 1984; Axelrod & Hamilton, 1981; Trivers, 1971), even when working indirectly through individuals' reputations (Nowak & Sigmund, 2005). When the long-run gains to cooperation eclipse the combined value of free riding and persistent mutual defection, reciprocity proves to be an effective mechanism of supporting cooperation. Empirical evidence, furthermore, indicates that reciprocity works over narrow timescales to induce cooperation: immediate retaliation for defection deters individuals from states of mutual defection and quick forgiveness of a defection (once a partner resumes cooperation) leads to sustained levels of mutual cooperation (S. S. Komorita, Hilty, & Parks, 1991). Moreover, the larger the number of entities employing a strategy of reciprocity, the greater the chances of bolstering mutual cooperation (S.S. Komorita, Parks, & Hulbert, 1992). Indeed, in an experiment in which Komorita et al. (1992, p.613) populated a group of 9 participants with 6 automated confederates that each reciprocated the choice of the most-cooperative of 3 non-confederate participants, the average proportion of cooperative plays equaled 0.47, whereas it equaled 0.24 in a condition in which the automated confederates randomly chose to cooperate with 0.33 probability. These results implied a percent increase in cooperation of 95.8%.

These findings indicate how meeting cooperation with cooperation can sustain pro-social interactions and such efforts can be employed in the COVID-19 response. Public officials should encourage members of the community to offer remote methods of assistance to each other (e.g., performing yard work for elderly neighbors so they can remain indoors) in order to stoke reciprocal cooperation. Also, thanking individuals for participating in videoconferences, as opposed to in-person meetings, draws attention to others' cooperative actions and rewards them for it. Withholding such support for those who defy cooperative response measures also should be encouraged. Executing such reciprocity in a systematic, large-scale intervention may be possible by repurposing tools for promoting other forms of cooperation. For instance, mass delivery of gratitude expressions for individuals who turn out in elections—a behavior that itself entails cooperation in a social dilemma (Fowler, 2006; Munger & Munger, 2015)—increases turnout in subsequent elections (Panagopoulos, 2011). Adapting such measures for public service announcements in which public officials thank individuals for engaging in social distancing could be done readily and it constitutes a cooperative reward as a response to initial cooperation. Doing so would spotlight the cooperation at the heart of social distancing and implement the reciprocity shown to generate cooperation in social dilemmas.

Research on reciprocity also suggests a potential means by which jurisdictions around the country can share resources cooperatively in order to combat COVID-19 transmission. Currently, large urban areas have encountered the most-pressing need for medical supplies and personnel, while rural areas have only just begun to experience cases. A reciprocal scheme of cooperatively sharing personnel and resources across communities could improve resource availability as the virus moves from locality to locality. Smaller



jurisdictions can offer personnel and resources to large urban areas experiencing early increases in cases, while large urban areas can pledge resources to smaller jurisdictions for treating future cases. Such schemes can involve immediate cash payments from urban areas with larger, more-flexible budgets to assure future cooperation. The literature on social dilemmas suggests that reciprocal cooperation of this variety has a strong probability of success and public officials could execute it to allocate resources in proportion to the virus' movement across jurisdictions.

*Lessons for public officials:*

1. Reward cooperative acts
2. Forge reciprocal exchange programs across communities hit by the pandemic at different times

### **Bolstering Mandatory Cooperative Measures via Altruistic Punishment**

Continued transmission of the coronavirus in high-density areas has resulted in aggressive measures to slow transmission. These measures include mandatory forms of cooperation such as sheltering in place and refraining from purchasing scarce medical supplies. The publicity associated with these mandatory cooperative measures raises the possibility that individuals will respond to non-cooperation through a well-studied, impulsive behavior that supports cooperation: costly punishment. Costly punishment (a.k.a. altruistic punishment) involves incurring a loss to reduce the welfare of free riders, which benefits the group (Fehr & Gächter, 2000; Fehr & Gächter, 2002; Yamagishi, 1986). This phenomenon comes in two varieties—decentralized and centralized costly punishment—both of which bear on public officials' responses to social dilemmas resulting from the COVID-19 crisis.

Decentralized costly punishment involves uncoordinated sanctioning of non-cooperators. Despite the personal costs associated with it and the lack of joint orchestration, this behavior quickly increases rates of cooperation (Fehr & Gächter, 2000; Fehr & Gächter, 2002). Moreover, it does so before anyone engages in it. Evidence from public goods games with the option of costly punishment shows that random assignment of the punishment mechanism spurs higher rates of cooperation from the outset of an experiment—that is, prior to when others even indulge the option to pay a cost to sanction free riders (Fehr & Gächter, 2000). The mere recognition that punishment is possible disciplines defectors (Fehr & Gächter, 2000). This fact appears surprising because one would expect individuals to free ride on the provision of punishment itself—after all, it requires individuals to assume costs that benefit the group (Yamagishi, 1986). However, the anger that results from both having one's cooperation exploited (Fehr & Gächter, 2002) and seeing the inequalities created by free riding (C. T. Dawes, Fowler, Johnson, McElreath, & Smirnov, 2007; Johnson, Dawes, Fowler, McElreath, & Smirnov, 2009) spurs individuals to punish. Though they incur a material cost, altruistic punishers receive the psychological benefit of seeing justice served.

The effect of punishment on cooperation is substantial. Across sessions of a foundational set of public goods game experiments, Fehr & Gächter (2000, p. 985) observed a 210.8% increase in mean cooperative contributions to the public good in all periods with opportunities for decentralized costly punishment versus all periods without such opportunities. In conditions without punishment, participants contributed, on average, 3.7 units of a 20-point endowment (18.5%) to the public good, whereas they contributed an average of 11.5 points of their 20-point endowment (57.5%) in conditions involving punishment—a 39-percentage-point increase. Subsequent research has explained how

decentralized altruistic punishment figures centrally in humanity's large-scale cooperation (Bowles & Gintis, 2011; Fischbacher & Fehr, 2003) and appears prevalent among individuals who participate in collective endeavors that exhibit the social dilemma's incentive structure (Smirnov, Dawes, Fowler, Johnson, & McElreath, 2010). The role of decentralized costly punishment in addressing the social dilemmas involved in the COVID-19 response, however, warrants careful consideration.

On the one hand, decentralized costly punishment already has figured beneficially into private responses to the COVID-19 crisis. In the second weekend of March 2020, reports surfaced of a pair of brothers who had bought-up 17,000 bottles of hand sanitizer with the intent of charging high prices for them on Amazon and eBay (Wire, 2020). Although they themselves would have gained by facilitating the high-priced sales of price gougers, those companies viewed the behavior as exploiting the mutually beneficial trading they sought to support (Palmer, 2020). At a personal cost, the companies cut-off the trades and showed price-gouging defectors that their exploitation would be met with punishment—in this instance, in the form of a garage full of unsold product and the cost of holding that inventory. By the end of the weekend in which reports surfaced, the punishment had drilled a new mode of behavior in the brothers as one of them donated all 17,700 bottles of hand sanitizer (Nicas, 2020).

On the other hand, instances of misbegotten vigilantism raise the specter of decentralized costly punishment gone wrong. For instance, at the end of March 2020, media reports surfaced concerning an incident in which a band of armed residents felled a tree to block one of their neighbors from leaving home (Srikanth, 2020). The band of vigilantes suspected the neighbor of having COVID-19 and they assumedly deemed their activity to be

a way to halt violation of quarantine measures. This instance recalls the well-known finding that anti-social motivations can hijack decentralized costly punishment (Herrmann, Thöni, & Gächter, 2008).

These dueling anecdotes highlight ways that public officials should contemplate and use decentralized costly punishment in the COVID-19 response. For one, the anecdotes—and particularly the troubling, latter one—indicate that public officials ought to regard studies of costly punishment as evidence that the prospect of vigilante justice emerging in the COVID-19 crisis is real. Laboratory experiments show that individuals will spend substantial resources to levy fines for suffering relatively minor costs from free riders; this willingness might become more pronounced if individuals view free riding as a potential agent for spreading a deadly virus. Officials, accordingly, should emphasize bans on vigilantism in their communications to the community. Or, if they wish to provide positive guidance (Wegner et al., 1987), then officials ought to harness individuals' willingness to engage in decentralized costly punishment by setting clear standards of cooperation and communicating acceptable ways of engaging in sanctioning. For instance, setting prohibitions on the number of scarce products that an individual can buy at retail outlets and allowing stores to ban customers found to violate those quotas might be an acceptable method to redirect the cooperation-enhancing zeal of decentralized costly punishment. If such sanctioning measures are necessary for supporting cooperation, then the literature on decentralized costly punishment suggests that public officials should find ways to compensate the costly punisher so that the behavior can remain sustainable despite the costs of punishment. Otherwise, the opportunity for cooperators to free ride on the costs of sanctioning will convince punishers to refrain from the activity (see Fowler, 2005).

However, a better means of productively focusing individuals' willingness to incur the costs of punishment might be by designing ways for community members to support *centralized* costly punishment. Centralized costly punishment involves the designation of a sanctioning authority that receives payment in order to sanction free riders (O'Gorman, Henrich, & Van Vugt, 2008; Boyd, Gintis, & Bowles, 2010; Baldassarri & Grossman, 2011). Research indicates that centralized sanctioning increases cooperative contributions to public goods by 16.6% in the midst of an experiment, even when the punisher is chosen at random, and it increases those contributions by a further 8.8% when the centralized punishment authority has been selected via a legitimate election (Baldassarri & Grossman, 2011, p.11024, fig.1). Public officials seeking to harness this mechanism for cooperation could develop methods through which individuals can contribute to the activities of legitimate, centralized sanctioning authorities such as the police. The creation or publicizing of accounts for the receipt of charitable contributions to local policing efforts constitutes one means by which public officials can channel support for centralized sanctioning institutions. Also, developing tip lines to allow community members to inform centralized sanctioning authorities about violations of COVID-19 response measures—such as violation of shelter-in-place mandates—offers a means of leveraging individuals' willingness to aid in the penalizing of free-riders.

The latter measure, on its face, sits uncomfortably with the values of liberal, capitalistic societies and costly punishment more generally ought to be employed cautiously as a method to increase cooperation. Research on altruistic punishment across the globe, however, does show that individuals support the reprimanding of non-cooperation, even at a cost, and they exhibit a greater tendency to do so in localities in which altruism is more

common (Henrich et al., 2006). Prior research indicates that altruistic, pro-sociality correlates with a community's market integration (Henrich et al., 2005), thus suggesting that even communities with strong expectations surrounding personal liberties may support the imposition of punishment, at a personal cost, to reduce non-cooperation. Used carefully and directed solely for the purpose of correcting violations of agreed-upon standards of cooperation, costly punishment may serve as an important mechanism for sustaining the values with which it seems to conflict. In that light, it may represent a potentially powerful tool to protect the cooperative efforts needed to slow coronavirus transmission.

*Lessons for public officials:*

1. Set clear standards for what constitutes cooperative behavior and acceptable costly punishment
2. Anticipate vigilantism and remind individuals of its illegality
3. Introduce opportunities for individuals to donate specifically to policing measures related to COVID-19 response
4. Create hotlines for informing authorities about violations of COVID-19 response measures

**Preparing for Subsequent Social Dilemmas Resulting from the Coronavirus**

The social dilemmas at the heart of slowing coronavirus transmission may be the first of several such dilemmas that emerge as the result of the present crisis. Thus, lessons about how to promote cooperation may prove vital in the months to come.

For instance, market data now indicate that gun sales have increased markedly in the U.S. over the past several weeks, thus creating the possibility of a household-level variant of

the “security dilemma” (Jervis, 1978). The security dilemma translates the social dilemma to the international stage; in the model, countries recognize they are best off bolstering their security through military investment when others do not do so, yet knowledge of this possibility causes other countries to act on the same logic, creating an arms race and frictions that could spark war. Albeit a low probability scenario, a heavily armed population with high unemployment and resource deprivation could find itself in a situation in which households devote excessive resources to home security and the possibility of preemptive violence becomes real. Invoking informal mechanisms of cooperation may offer a social glue to ensure that individuals adhere to top-down, government policies that already regulate such situations.

The potential for prolonged transmission of the virus also raises a social dilemma that intersects with economic policy. Successful economic policy will spur consumption, which creates an incentive for businesses to return to normal operations and forestall measures of social distancing, even if those measures remain needed. All enterprises would be best off to cooperate and reduce transmission, but any enterprise doing so in a competitive market would run the risk of losing market share while a competitor expands operations. In turn, renewed outbreak of the virus might occur, leaving all enterprises with sick employees and an outcome that would have been worse than mutually agreeing to practices that would maintain lower levels of transmission. Using the mechanisms discussed above to prevent such problems may prove to be necessary.

The historical frequency of community challenges that resemble the social dilemma—from environmental preservation (Hardin, 1968) to online trading (Johnson & Smirnov, 2013)—suggest that the coronavirus crisis might generate further social dilemmas.

Some of these social dilemmas will be repeated games, such as the opportunities for intergovernmental cooperation that we discussed earlier in this article. Others, such as how individuals conduct themselves on public transit in large cities where encounters with others are fleeting, will be one-shot games. Gaining practice with mechanisms shown to support cooperation in the social dilemma may help communities confront these diverse challenges that loom on the horizon.

### **Conclusion**

Just as medical research can contribute to the development of a vaccine, research on the social dilemma can provide the guidance that public officials need to solve the social challenges associated with slowing transmission of COVID-19. Prior to widespread transmission, government needs to encourage individuals to trust their gut instincts in their social response to the disease, thus channeling individuals' instinctive tendency to cooperate. They also need to insulate pockets of cooperators from defectors. As officials begin to recommend non-mandatory measures to control virus spread, communities should promote common group identities, as well as strategies of reciprocity in which they highlight and reward cooperation while penalizing defection in soft ways that seek to restore cooperation. As more-stringent standards appear, private and public enterprises should draw on and support individuals' acceptance of the costs of punishing free riders, while admonishing vigilantism that might result from individuals' eagerness to discipline free riders.

Ultimately, the research underlying these findings has shown that humanity's resolution of social dilemmas has led to it flourishing (Bowles & Gintis, 2011; Fischbacher & Fehr, 2003; Nowak & Highfield, 2011). Now, if heeded, that research stands poised to help



humanity continue on that trajectory by guiding communities on how to slow transmission of COVID-19.

Author Version

## References

- Acker, L. (2020). Tillamook mayor to tourists: 'Pack your bags and leave immediately'. Oregonlive.com(March 22). doi: <https://www.oregonlive.com/coronavirus/2020/03/tillamook-mayor-to-tourists-pack-your-bags-and-leave-immediately.html>
- Aktipis, C. A. (2004). Know when to walk away: contingent movement and the evolution of cooperation. *Journal of theoretical biology*, 231(2), 249-260. doi: <https://doi.org/10.1016/j.jtbi.2004.06.020>
- Axelrod, R. (1984). *The Evolution of Cooperation*. New York, NY: Basic Books.
- Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211(27), 1390-1396.
- Baldassarri, D., & Grossman, G. (2011). Centralized sanctioning and legitimate authority promote cooperation in humans. *Proceedings of the National Academy of Sciences*, 108(27), 11023-11027. doi:10.1073/pnas.1105456108
- Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature*, 442, 912-915. doi: 10.1038/nature04981
- Bowles, S., & Gintis, H. (2011). *A Cooperative Species: Human Reciprocity and Its Evolution*. Princeton: Princeton University Press.
- Boyd, R., Gintis, H., & Bowles, S. (2010). Coordinated Punishment of Defectors Sustains Cooperation and Can Proliferate When Rare. *Science*, 328, 617-620.
- Choi, J.-K., & Bowles, S. (2007). The Coevolution of Parochial Altruism and War. *Science*, 318(5850), 636-640. doi:10.1126/science.1144237
- Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007). Egalitarian motives in humans. *Nature*, 446(7137), 794-796.
- Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, 31, 169-193.
- Dawes, R. M., Van De Kragt, A. J. C., & Orbell, J. M. (1988). Not me or thee but we: The importance of group identity in eliciting cooperation in dilemma situations: Experimental manipulations. *Acta Psychologica*, 68(1-3), 83-97. [https://doi.org/10.1016/0001-6918\(88\)90047-9](https://doi.org/10.1016/0001-6918(88)90047-9)

- Dawson, A. (2020). UFC president Dana White says he's spoken to Trump, doesn't think there's any point panicking about coronavirus, and vows UFC events will continue. INSIDER. Retrieved from <https://www.insider.com/dana-white-dont-panic-ufc-will-continue-to-hold-events-2020-3>
- De Herdt, T. (2003). Cooperation and Fairness: The Flood—Dresher Experiment Revisited. *Review of Social Economy*, 61(2), 183-210. Retrieved from [www.jstor.org/stable/29770203](http://www.jstor.org/stable/29770203)
- Fehr, E., & Gächter, S. (2000). Cooperation and Punishment in Public Goods Experiments. *American Economic Review*, 90(4), 980-994. doi:10.1257/aer.90.4.980
- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415, 137-140.
- Fischbacher, U., & Fehr, E. (2003). The nature of human altruism. *Nature*, 425(6960), 785-791.
- Fowler, J. H. (2005). Altruistic punishment and the origin of cooperation. *Proceedings of the National Academy of Sciences USA*, 102(19), 7047-7049.
- Fowler, J. H. (2006). Altruism and Turnout. *The Journal of Politics*, 68(3), 674-683. doi:10.1111/j.1468-2508.2006.00453.x
- Givas, N. (2020). Washington state health officials give coronavirus guidance: 'Don't panic, get prepared'. Fox News. Retrieved from <https://www.foxnews.com/health/coronavirus-washington-state-update>
- Greenstone, M., & Nigam, V. (2020). Does Social Distancing Matter? Becker Friedman Institute for Economics Working Paper No. 2020-26. <http://dx.doi.org/10.2139/ssrn.3561244>
- Gutiérrez-Roig, M., Gracia-Lázaro, C., Perelló, J., Moreno, Y., & Sánchez, A. (2014). Transition from reciprocal cooperation to persistent behaviour in social dilemmas at the end of adolescence. *Nature Communications*, 5(1), 4362. doi:10.1038/ncomms5362
- Hardin, G. (1968). The Tragedy of the Commons. *Science*, 162(3859), 1243-1248.

- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., . . . Tracer, D. (2005). "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, 28(6), 795-815. doi:10.1017/S0140525X05000142
- Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., . . . Ziker, J. (2006). Costly Punishment Across Human Societies. *Science*, 312(5781), 1767-1770. doi:10.1126/science.1127333
- Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial Punishment Across Societies. *Science*, 319(5868), 1362-1367. doi:10.1126/science.1153808
- Hoffower, H. (2020). Travelers worldwide are cancelling trips over coronavirus fears — but millennials are snapping up cheap flights despite the risks, and it says a lot about who they are as a generation. *Business Insider*. Retrieved from <https://www.businessinsider.com/millennials-buying-cheap-coronavirus-flights-2020-3>
- Holshue, M. L., DeBolt, C., Lindquist, S., Lofy, K. H., Wiesman, J., Bruce, H., . . . Pillai, S. K. (2020). First Case of 2019 Novel Coronavirus in the United States. *New England Journal of Medicine*, 382(10), 929-936. doi:10.1056/NEJMoa2001191
- Isler, O., Maule, J., & Starmer, C. (2018). Is intuition really cooperative? Improved tests support the social heuristics hypothesis. *PLoS ONE*, 13(1), e0190560. doi:10.1371/journal.pone.0190560
- Jagau, S., & van Veelen, M. (2017). A general evolutionary framework for the role of intuition and deliberation in cooperation. 1, 0152. doi:10.1038/s41562-017-0152 <https://www.nature.com/articles/s41562-017-0152#supplementary-information>
- Jervis, R. (1978). Cooperation under the Security Dilemma. *World Politics*, 30(2), 167-214. doi: <https://doi.org/10.2307/2009958>
- Johnson, T., Dawes, C. T., Fowler, J. H., McElreath, R., & Smirnov, O. (2009). The role of egalitarian motives in altruistic punishment. *Economics Letters*, 102(3), 192-194.
- Johnson, T., & Smirnov, O. (2013). Cooperate with equals: A simple heuristic for social exchange. In R. Hertwig, U. Hoffrage, & A. B. C. Research Group (Eds.), *Simple Heuristics in a Social World* Oxford University Press.

- Komorita, S. S., Hilty, J. A., & Parks, C. D. (1991). Reciprocity and Cooperation in Social Dilemmas. *Journal of Conflict Resolution*, 35(3), 494-518.  
doi:10.1177/0022002791035003005
- Komorita, S. S., Parks, C. D., & Hulbert, L. G. (1992). Reciprocity and the Induction of Cooperation in Social Dilemmas. *Journal of Personality and Social Psychology*, 62(4), 607-717.
- Krajbich, I., Bartling, B., Hare, T., & Fehr, E. (2015). Rethinking fast and slow based on a critique of reaction-time reverse inference. *Nature Communications*, 6(1), 7455.  
doi:10.1038/ncomms8455
- Lunn, P.D., Belton, C.A., Lavin, C., McGowan, F.P., Timmons, S., & Robertson, D.A. (2020). Using behavioral science to help fight the coronavirus. *Journal of Behavioral Public Administration*, 3(1), 1-35. doi: 10.30636/jbpa.31.147
- Melamed, Samantha. (2020). Coronavirus fears ignite a civil war as families, friends, neighbors class over what's safe. *Philadelphia Inquirer*. Retrieved from <https://www.inquirer.com/health/coronavirus/coronavirus-covid-pandemic-philadelphia-montgomery-county-pennsylvania-families-20200316.html>
- Munger, M.C., & Munger, K.M. (2015). Voting as a Collective Action Problem.. In M. C. Munger (Ed.), *Choosing in Groups: Analytical Politics Revisited* (pp. 175-190). Cambridge: Cambridge University Press.
- Nicas, J. (2020). The Man With 17,700 Bottles of Hand Sanitizer Just Donated Them. *The New York Times*. Retrieved from <https://www.nytimes.com/2020/03/15/technology/matt-colvin-hand-sanitizer-donation.html>
- Nowak, M. A. (2006a). *Evolutionary Dynamics: Exploring the Equations of Life*. Cambridge, MA: Belknap Press.
- Nowak, M. A. (2006b). Five rules for the evolution of cooperation. *Science*, 314(5805), 1560-1563.
- Nowak, M. A., & Highfield, R. (2011). *Supercooperators*. New York: Free Press.
- Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, 437(7063), 1291-1298.

Press Office of New York City Hall. (2020). Mayor de Blasio Issues New Guidance to New Yorkers. New York City Hall Press Release(March 20, 2020).

doi:<https://www1.nyc.gov/office-of-the-mayor/news/173-20/mayor-de-blasio-issues-new-guidance-new-yorkers>

O'Gorman, R., Henrich, J., & Van Vugt, M. (2009). Constraining free riding in public goods games: Designated solitary punishers can sustain human cooperation. *Proceedings of the Royal Society B Biological Sciences*, 276, 323–329.

Ohtsuki, H., Hauert, C., Lieberman, E., & Nowak, M. A. (2006). A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441(7092), 502-505. doi:[http://www.nature.com/nature/journal/v441/n7092/supinfo/nature04605\\_S1.html](http://www.nature.com/nature/journal/v441/n7092/supinfo/nature04605_S1.html)

Palmer, A. (2020). Amazon tells sellers it's not taking new listings for face masks or hand sanitizer amid coronavirus price gouging. CNBC. Retrieved from <https://wgntv.com/news/brothers-buy-17000-bottles-of-hand-sanitizer-before-amazon-pulled-sales/>

Panagopoulos, C. (2011). Thank you for Voting: Gratitude Expression and Voter Mobilization. *Journal of Politics*, 73(3), 707-717.

Pancevski, B., & Meichtry, S. (2020). A Generational War Is Brewing Over Coronavirus. *Wall Street Journal*. Retrieved from <https://www.wsj.com/articles/a-generational-war-is-brewing-over-coronavirus-11584437401>

Poundstone, W. (1992). *Prisoner's Dilemma: John von Neumann, Game Theory and the Puzzle of the Bomb*. New York: Doubleday.

Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, 489, 427-430. <https://doi.org/10.1038/nature11467>

Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., & Greene, J. D. (2014). Social heuristics shape intuitive cooperation. *Nature Communications*, 5, 3677.

Rogers, T., & Marshall, A. (2020). Berlin Parties On, Despite the Coronavirus. *The New York Times*. Retrieved from <https://www.nytimes.com/2020/03/12/arts/music/berlin-clubs-coronavirus.html>

- Ruoran, L., Rivers, C., Tan, Q., Murray, M. B., Toner, E., & Lipsitch, M. (2020). The Demand for Inpatient and ICU Beds for COVID-19 in the US: Lessons From Chinese Cities. Harvard University DASH Repository. doi: <http://nrs.harvard.edu/urn-3:HUL.InstRepos:42599304>
- Santos, F. C., & Pacheco, J. M. (2005). Scale-Free Networks Provide a Unifying Framework for the Emergence of Cooperation. *Physical Review Letters*, 95, 098104.
- Smirnov, O., Dawes, C. T., Fowler, J. H., Johnson, T., & McElreath, R. (2010). The Behavioral Logic of Collective Action: Partisans Cooperate and Punish More Than Nonpartisans. *Political Psychology*, 31(4), 595-616. doi:10.1111/j.1467-9221.2010.00768.x
- Superville, D., & Miller, Z. (2020). Trump, on 1st death from virus in US: 'No reason to panic'. AP News Wire. Retrieved from <https://apnews.com/ad97746163f019c576837838f8a6d267>
- Srikanth, Anagha. (2020). Armed neighbors cut down a tree to block man's driveway because they thought he had coronavirus. *The Hill*. Retrieved from <https://thehill.com/changing-america/respect/diversity-inclusion/490475-a-man-reports-armed-neighbors-cut-down-a-tree>
- Szabó, G., & Fath, G. (2007). Evolutionary games on graphs. *Physics reports*, 446(4), 97-216.
- Taylor, P. D., & Jonker, L. B. (1978). Evolutionary stable strategies and game dynamics. *Mathematical biosciences*, 40(1-2), 145-156.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46(1), 35-57.
- Van Bavel, J., Boggio, P., Capraro, V., Cichocka, A., Cikara, M., Crockett, M., . . . Willer, R. (2020). Using social and behavioural science to support COVID-19 pandemic response. *PsyArXiv Preprints*. doi:10.31234/osf.io/y38m9
- Wire, N. M. (2020). Brothers Buy 17,000 Bottles of Hand Sanitizer Before Amazon Pulled Sales. *WGN9 Chicago*. Retrieved from <https://wgntv.com/news/brothers-buy-17000-bottles-of-hand-sanitizer-before-amazon-pulled-sales/>

Yamagishi, T. (1986). The provision of a sanctioning system as a public good *Journal of Personality and Social Psychology*, 51(1), 110–116. <https://doi.org/10.1037/0022-3514.51.1.110>

Yamagishi, T., & Mifune, N. (2016). Parochial altruism: does it explain modern human group psychology. *Current Opinion in Psychology*, 7(February), 39-43. <https://doi.org/10.1016/j.copsyc.2015.07.015>

Zetlin, M. (2020). Here's How to Host a Virtual Happy Hour and Why It Will Improve Your Mood. Inc.(March 20). doi:<https://www.inc.com/minda-zetlin/virtual-happy-hour-cocktails-video-chat-working-from-home.html>

<sup>1</sup> The public goods game presents a multi-person, continuous version of the social dilemma. In the public goods game, each individual  $i$  in a group of  $n$  individuals can contribute some portion of an endowment,  $w_i$ , to a common pool and retain the non-contributed portion of the endowment. The total contributions  $\sum_{i=1}^n w_i$  are multiplied by  $m$  and, then, divided evenly across participants. When  $1 < m < n$ , participants have an incentive to withhold contributions in the hopes of free riding on other participants' provisioning of the common pool.

<sup>2</sup> We thank an anonymous reviewer both for pointing out this additional reason that public officials should avoid cautioning against behaviors stemming from panic and for informing us of Wegner et al. (1987), which justifies that reason.

<sup>3</sup> At the time, Flood and Dresher used the term 'prisoner's dilemma' to describe the dilemma, due to the compelling substantive motivation that one of their colleagues used to depict the scenario (Poundstone, 1992). Like two prisoners in separate cells who the police have invited to snitch on their accomplice, the participants in the prisoner's dilemma are jointly best off to uphold the Code of Omerta; however, the temptation to narc for individual gain leads both to do so and to suffer the sub-optimal outcome of a severe sentence. The label 'prisoner's dilemma' continues to be used to refer to a 2-person game in which individuals have a choice between two discrete options, Cooperate (C) or Defect (D). Both parties cooperating results in  $R$ , but individuals earn the best outcome,  $T$ , when they defect on cooperation and impose the penalty  $S$  on cooperators. Since all know that possibility, players choose to defect to exploit others or avoid exploitation, producing the outcome  $P$ , which accrues to mutual defection. The payoff ordering  $T > R > P > S$  describes the incentives leading to that outcome.