Cooperation in Social Dilemmas
28 Years of Research on Collective Goods

*Svenja Hippel & Dorothee Mischkowski**

"When God expelled Adam and Eve from paradise, he created rationality. Not only would humans have to cope with scarcity, and hence have to trade some desired goods for others. They would even be pitted against each other. Individually rational behaviour would lead to social dilemmas." (Engel, 2004, p. 2)

A. Introduction

Cooperation, understood as the willingness to forgo immediate self-interest for the sake of collective benefit, forms a cornerstone of functioning societies. Yet in social dilemmas, where individual payoff maximization conflicts with group welfare, cooperation becomes both crucial and inherently puzzling. Christoph Engel has fundamentally advanced our understanding of this tension through a genuinely interdisciplinary research program that unites law, economics, and experimental psychology. His work shows how institutional structures – legal rules, social norms, and economic incentives – shape the conditions under which cooperation can emerge or fail, and how psychological factors such as expectations about others and individual differences in social preferences guide behavior in these environments. The resulting research represents a level of cross-disciplinary synthesis that remains rare in the academic landscape.

Understanding the mechanisms that foster or hinder cooperation has far-reaching implications. Social dilemmas underlie a wide range of real-world challenges. "In fact, people's selfishness is often the biggest obstacle to reaching better social outcomes. From environmental pollution over tax evasion, corruption and misappropriation, to doping, queuing, and bank runs, the list of situations is sheer endless in which the collective interests of society and the selfish interests of its individual members are

* Both authors contributed equally to this article.

at odds" (Engel & Kurschilgen, 2015, p. 2). Even scientific integrity is not immune to adverse dilemma structures: Engel (2015) discussed various mechanisms for ensuring scientific integrity nonetheless (e.g., sanctions and internalized norms) to prevent personal incentives from dominating, showing, however, that none of these solutions are perfect. Adding further complexity, there are situations such as competition in markets, where social dilemmas might even be a socially desirable feature (Engel, 2009; Engel & Zhurakhovska, 2011).

Earlier research already identified key determinants of human cooperation (e.g., Camera & Casari, 2009; Fischbacher & Gächter, 2010; Friedman & Oprea, 2012). Yet important questions remained about why cooperation sometimes thrives despite misaligned incentives – and why, in other instances, it ultimately collapses. The work of Christoph Engel and his group at the Max Planck Institute for Research on Collective Goods has provided critical new insights into these dynamics.

B. Determinants of Cooperation

In what follows, we organize our contribution around the key determinants of cooperation that Christoph Engel has investigated. We begin with contextual factors – such as institutional mechanisms (e.g., sanctions and punishment), the role of authorities, social norms, and negative externalities – that shape the extent of cooperative behavior. A second section highlights individual differences and the dynamics of conditional cooperation, with particular attention to the fragility of cooperation following negative experiences (e.g., exploitation by interaction partners). Finally, we turn to the psychological reactions to insufficient cooperation, including Christoph Engel's pioneering work on deliberate ignorance.

I. Sanctions and Punishment

Peer punishment provides contributors to public goods with an additional tool for enforcing cooperation within the group. Fehr and Gächter (2000) demonstrated that individuals are indeed willing to punish defectors despite bearing personal costs, and that such punishment can help sustain cooperation.

In Engel (2014), the central puzzle is why even weak sanctions – those too mild to deter a payoff-maximizing free rider – nonetheless stabilize cooperation. Here, Christoph Engel integrates formal theorizing with experimental evidence, demonstrating that the solution to this puzzle is rooted in social preferences – specifically, advantageous inequity aversion. A parsimonious model illustrates how even non-deterrent sanctions can amplify weak inequity aversion, shifting the incentive balance in favor of cooperation. Experimentally, participants increased their contributions following punishment, even when sanctions were too small to deter a strictly payoff-maximizing individual. This response was markedly stronger among participants with higher inequity aversion, indicating that the stabilizing effect of sanctions depends on social preferences, not deterrence alone. Together, these findings reveal that imperfect sanctions complement underlying distributive concerns, relying on preference-based cooperation motives alongside – not instead of – material incentives.

Engel, Kube, and Kurschilgen (2021) turn to the effect of selective information on cooperation in the presence of sanctions. In their experiments, subjects received favorable or unfavorable histories of cooperation from unrelated groups before playing public good games with and without peer- and counter-punishment. Unfavorable information sharply lowered initial contributions through influencing initial beliefs and substantially reduced the effectiveness of punishment, preventing groups from recovering once cooperation started at a low level. Favorable information, by contrast, helped stabilize contributions in less cooperative environments but had little effect where cooperation was already high. Thus, while punishment can in principle enforce norms, its effectiveness is fragile and highly dependent on the informational environment in which expectations are formed.

II. Authorities

While peer punishment allows contributors to stabilize cooperation on their own, an authority – a person with elevated status and decision-making power – may also be present, shaping the situation through their actions. Such an authority figure may be either a member of the group or an external third party.

Engel and Hippel (2017) demonstrate both the promise and limits of such an institutionalized authority in solving social dilemmas. They bridge mechanism design and experimental public good research by studying the behavior of experimental social planners. In their design, a randomly assigned authority could impose a contribution rule on group members with heterogeneous valuations for a public good. Across treatments, the authors varied whether valuations were common knowledge or private information (with scope for misrepresentation), and whether the authority was personally affected by the rule.

Disinterested planners predominantly opted for rules equalizing final payoffs across participants, even when efficiency would have dictated otherwise. Strikingly, they were overly optimistic about truth telling: Most expected active participants to report valuations honestly, despite strong incentives to misrepresent. Once the planner's own interests came into play, optimism gave way to self-serving behavior. When deciding under a veil of ignorance, planners balanced efficiency and fairness, but they still overestimated others' willingness to reveal valuations truthfully. When they knew their own valuation, decisions became straightforwardly selfish. While many experimental planners acted in line with fairness norms when being personally unaffected, the combination of optimism about others' honesty and the temptation of self-interest meant that even institutional solutions remained precarious. This cautions against assuming that the mere presence of a (hopefully) benevolent authority guarantees efficient public good provision.

III. Externalities and Voting

While punishment expands the players' action space by adding behavioral response options, authorities are introduced as additional actors who constrain or enforce the boundaries of choice. However, public good provision is typically embedded within a broader social context. Externalities –

positive or negative effects on bystanders – are a defining and ubiquitous feature of such environments and may meaningfully shape contributions to public goods.

Engel and Rockenbach (2009) examine in an initial study how externalities influence cooperation when passive outsiders are affected. They find that the presence of either negative or positive externalities in a group's public good provision had almost no impact on contribution levels, provided that no explicit provision norms exist. This was the case in the first and last phase of their three-phase experiment. However, when norms emerged endogenously – implemented through a non-binding vote on the desired contribution level before the beginning of the second phase – negative externalities became internalized and contributions declined. Somewhat surprisingly, positive externalities also reduced provision. The authors attribute this to conditional cooperators who face a twofold risk: Earning lower payoffs not only than free riders within the group but also than passive outsiders. Additionally, there was no spillover of normativity after the contribution norm was abolished again in the last phase. It continued to only matter indirectly since it has induced some favorable experiences.

The experimental design of Engel and Rockenbach (2009) was subsequently refined to investigate two aspects in greater depth. Both follow-up studies continued to examine the effect of externalities on cooperation, but they do so from different angles. Engel and Zhurakhovska (2014) focused on the apparent lack of restraint in imposing harm on passive bystanders, whereas Engel and Rockenbach (2014) took a closer look at the role of giving people a voice through voting.

Engel and Zhurakhovska (2014) simplified the original design to clarify what drives conditional cooperation under negative externalities. A passive third party was harmed whenever at least one of two active players cooperated. By varying the intensity of this harm and eliciting participants' expectations about others' cooperation, they separated cognitive from motivational influences. Their results show that cooperation declined as outsider harm intensified, and this decline could not be explained by inequity aversion or reciprocity. Instead, guilt aversion was the central mechanism: Insiders experience discomfort from jointly harming outsiders, reducing their willingness to cooperate. The findings also suggest that greater strategic uncertainty heightens insiders' sensitivity to harming outsiders.

Engel and Rockenbach (2014) use a design similar to the first two

phases of Engel and Rockenbach (2009), but add a control treatment without the possibility of voting. This allows them to distinguish voting effects from mere restart effects. Their design enables them to study to what extent giving people a voice helps resolve the normative conflict that arises when a social dilemma is embedded in a broader social context. The results illustrate both the potential and the limits of such institutional voice. When no externalities were present, voting raised cooperation, showing that even a mere recommendation could coordinate insiders. When externalities were negative, voting lead to lower contributions, thereby successfully internalizing outsider harm. Yet when externalities were positive, voting failed to produce higher contributions. Although groups overwhelmingly voted for high provision, insiders complied only weakly, largely due to adverse payoff comparisons similar to those found in Engel and Rockenbach (2009): Group members were reluctant to contribute when outsiders received windfall gains.

Taken together, these findings on the role of passive outsiders underscore the importance of institutional environments that protect contributors from being exploited, particularly when public-good provision generates positive externalities that allow outsiders to benefit more than insiders. Conversely, when policymakers aim to protect global rather than local public goods, heavy-handed intervention may not be required to shield those harmed by negative externalities: Guilt aversion already helps minimizing harmful spillovers. However, in larger groups, institutions that give outsiders a voice and enable coordination on lower contribution rates can help reduce harmful effects.

IV. Nudges

Nudges do not formally alter the actions or consequences available to contributors, but are interventions that aim to steer behavior in a particular direction while preserving freedom of choice (Hertwig & Grüne-Yanoff, 2017). Therefore, policymakers and researchers have been drawn to exploit their potential. Although critiques and caveats have accumulated over time, nudges remain a widely studied tool because they may offer, among other benefits, a way to encourage contributions to public goods without requiring drastic institutional intervention.

Engel and Kurschilgen (2020) investigate whether asking individuals to articulate their own normative goals can help increase contributions in

social dilemmas. In their experiments, they contrast two framings of this nudge: Participants were asked either what every group member should optimally contribute or what they should contribute at minimum. In their first study, the "minimum" framing significantly increased cooperation and delayed the typical decline of cooperation across rounds. This effect was traced to a dual mechanism: self-set standards shaped participants' own willingness to contribute (a self-image channel) and influenced their expectations about others' choices (an indirect belief channel).

However, two later replication studies reveal the limits of articulating own normative goals. With nearly identical designs, but in contexts where baseline cooperation was already higher, neither the "minimum" nor the "optimal" framing produced a measurable effect on contributions. Importantly, the psychological process remained intact and stated norms still predicted individual behavior.

Engel and Kurschilgen (2020) thus caution that nudges built on self-set normative standards are highly context dependent: They may work when cooperative behavior is fragile, but lose traction when norms of cooperation are already entrenched. Their findings nuance the optimism surrounding behavioral policy tools, showing that although self-image can be a powerful motivator, the effectiveness of such nudges depends critically on whether the social dilemma was sufficiently pronounced prior to the intervention.

V. Heterogeneity in Cooperativeness: Individual Differences and Underlying Motives

Even in similar environments, cooperation is rarely uniform: Individuals differ systematically in their cooperativeness. A research line by Christoph Engel and colleagues has provided key insights into the dispositional foundations and underlying motives of cooperation behavior, highlighting how cooperation arises from the psychologically meaningful interplay of situational cues and individual characteristics.

A particularly dominant motive underlying (lacking) cooperation is the aversion to being taken advantage of: People are willing to cooperate, but conditional on mutual cooperation. Negative experiences, in turn, can trigger a disproportionate decline in cooperation (Engel & Rockenbach, 2024). Engel and Zhurakhovska (2014) argue that, based on defining environmental features of the prisoner's dilemma – most notably the risk of

exploitation alongside the possibility of maximizing joint gains – cooperation is best understood as the joint outcome of multiple motives, not of a single one alone. By taking efficiency concerns, conditional cooperation, fear, and greed into account, they show that these motives predict cooperation most reliably when estimated together, as some only become predictive when controlling for others. Through this lens, the prisoner's dilemma reflects a dynamic combination of (pro-)social preferences and interpersonal risk calculus that explain when cooperation becomes, for the individual, worth taking.

Importantly, these motives help explain not only when people choose to cooperate in single (i.e., one-shot) interactions, but more importantly so across repeated interactions. Engel and Rockenbach (2024) revisit one of the most robust findings in experimental economics: The gradual decline of contributions in repeated public good games. Building on Fischbacher and Gächter's (2010) influential claim that this decay reflects an inherent imperfection in conditional cooperation, they reanalyze the original data and introduce a more nuanced interpretation. Classifying participants into selfish free-riders (short- or far-sighted) and conditional cooperators, they show that the latter group actually tracked its own beliefs about others' contributions remarkably closely. The apparent imperfection was largely attributable to free-riders. Conditional cooperators only deviated from their beliefs after experiencing disappointment – when others contributed less than expected.

A series of simulations confirmed that disappointment asymmetry, rather than inherent imperfection, explains much of the observed decline. Conditional cooperators adjusted more strongly to negative than to positive experiences, and their behavior was further destabilized when even a single free-rider was present in the group. The analysis implies that cooperation is precarious not because conditional cooperators are insincere, but because they are vulnerable to the corrosive impact of free-riding and noisy behavior. Policy implications follow: Broad-based nudges may be less effective than targeted interventions aimed at preventing free-riding or mitigating its visibility. By highlighting the asymmetry of reactions, Engel and Rockenbach refine our understanding of conditional cooperation and its limits, and underscore how fragile collective action becomes once negative experiences enter the social learning process.

Desmet and Engel (2021) follow up on the motivation to conditionally cooperate in the context of rule compliance: They examine how compliance is influenced by the behavior of others. Using arbitrary, unen-

forced upper bounds in a real-effort task, they show that rule compliance became stronger when group members adhered to the rule, creating a robust form of behavioral dependence that resembles conditional cooperation. Unlike cooperation behavior in social dilemmas, however, conditional rule adherence appeared to be buffered against breakdowns, as each additional rule-abiding group member reinforced rule adherence. Conditionality was strongest among marginal rule followers, whose compliance was most strongly promoted by positive social information when their basic propensity to follow rules was low.

The findings highlight a deontological component in cooperative behavior: Individuals consider compliance to be inherently appropriate, but express this more consistently in their behavior when others also adhere to the rules. This dynamic complements multi-motive models of cooperation by indicating that rule following draws on motives similar to those underlying cooperation, albeit in a different and partly attenuated manner.

VI. Low-cost Cooperation and Deliberate Ignorance

People are generally willing to share parts of their resources (Engel, 2011), but cooperation particularly flourishes when costs are low, as shown in research on social mindfulness and everyday prosocial behavior (Van Doesum et al., 2025).

Engel and Van Lange (2021), for instance, demonstrated that people are generally socially mindful – that is, considerate of others' choice options and outcomes – when the costs are minimal. Yet even small increases in cost lead to a rapid decline in socially mindful behavior. Importantly, cost sensitivity already becomes visible in patterns of (reduced) information search and attention allocation: Offer et al. (2024) found that a substantial share of participants deliberately ignored information about unfairness when doing so spared them the burden of inflicting costly punishment on norm transgressors. This phenomenon illustrates a broader principle: Deliberate ignorance can serve as a functional strategy that protects individuals from motivational conflict, cognitive overload, or normative obligations (Hertwig & Engel, 2016). In this sense, choosing *not* to know can be understood as a form of resource management – even though actively seeking and integrating information is only minimally costly (Engel & Hertwig, 2021). Moreover, deliberate ignorance is not confined to the individual level. Institutional and legal settings often create or regulate

opportunities for ignorance to balance competing demands of efficiency, accountability, and fairness, thereby shaping when knowledge is required and when ignorance is tolerated or even promoted (Teichman et al., 2021).

Taken together, these findings indicate that the costliness of action shapes cooperation behavior itself, while deliberate ignorance can serve as a way to avoid the additional costs that may arise from responding to missing or insufficient cooperation – for instance, the psychological and material burdens of punishing norm violations. In this sense, choosing not to know can function as a protective strategy in the face of potentially costly reactions to others' non-cooperation.

C. Conclusion and Outlook

Across nearly three decades of research, Christoph Engel has shown that cooperation reflects a finely balanced, sensitive interplay of psychological motives and situational cues. And while no single mechanism can fully solve the problem of collective action, Christoph Engel's research has profoundly refined our understanding of when cooperation is possible, when it fails, and which levers we can use to strengthen the common good.

Naturally, the question remains whether the puzzle of cooperation will ever be completely solved. Earlier in his career – then still deeply immersed in the challenges of waste management – Christoph Engel addressed a very tangible social dilemma: He demonstrated that even something as mundane as separating household waste is not merely a technical matter, but a fundamentally social one. With hindsight, this early work reads almost like a prologue to his later insights: Even the most elegant system works only when individuals understand, expect, and trust that others will do their part. From the garbage bin at the curb to the repeated public-good game in the lab, a clear thread connects Christoph Engel's work: A persistent curiosity about what makes humans cooperative beings – and what gets in their way.

More recently, this curiosity has turned toward a new domain: Large language models, whose potential contributions include advancing our understanding of the persistent under-provision of public goods. If, one day, a large language model manages to solve this problem, it will almost certainly have been informed and guided by Christoph Engel's work.

References

Camera, G., & Casari, M. (2009). Cooperation among strangers under the shadow of the future. A*merican Economic Review*, *99*(3), 979–1005. https://doi.org/10.1257/aer.99.3.979.

Desmet, P. T. M., & Engel, C. (2021). People are conditional rule followers. *Journal of Economic Psychology, 85*, 102384. https://doi.org/10.1016/j.joep.2021.102384.

Engel, C. (2004). Social dilemmas, revisited from a heuristics perspective. *Preprints of the Max Planck Institute for Research on Collective Goods,* No. 2004/4. https://dx.doi.org/10.2139/ssrn.539442.

Engel, C. (2009). Competition as a socially desirable dilemma – Theory vs. experimental evidence. *Preprints of the Max Planck Institute for Research on Collective Goods,* No. 2009/24. http://dx.doi.org/10.2139/ssrn.1444270.

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics, 14*(4), 583–610. https://doi.org/10.1007/s10683-011-9283-7.

Engel, C. (2014). Social preferences can make imperfect sanctions work: Evidence from a public good experiment. *Journal of Economic Behavior & Organization, 108*, 343–353. https://doi.org/10.1016/j.jebo.2014.02.015.

Engel, C. (2015). Scientific disintegrity as a public bad. *Perspectives on Psychological Science, 10*(3), 361–379. https://doi.org/10.1177/1745691615577865.

Engel, C., & Hertwig, R. (2021). Deliberate ignorance: Present and future. In R. Hertwig & C. Engel (Eds.), *Deliberate ignorance: Choosing not to know* (pp. 317–332). Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/13757.003.0004.

Engel, C., & Hippel, S. (2017). Experimental social planners: Good natured, but overly optimistic. *Preprints of the Max Planck Institute for Research on Collective Goods*, No. 2017/23. https://dx.doi.org/10.2139/ssrn.3074350.

Engel, C., Kube, S., & Kurschilgen, M. (2021). Managing expectations: How selective information affects cooperation and punishment in social dilemma games. *Journal of Economic Behavior & Organization, 187*, 111–136. https://doi.org/10.1016/j.jebo.2021.04.029.

Engel, C., & Kurschilgen, M. (2015). The jurisdiction of the man within: Introspection, identity, and cooperation in a public good experiment. *Preprints of the Max Planck Institute for Research on Collective Goods*, No. 2015/1. https://dx.doi.org/10.2139/ssrn.2544579.

Engel, C., & Kurschilgen, M. (2020). The fragility of a nudge: The power of self-set norms to contain a social dilemma. *Journal of Economic Psychology, 81*, 102293. https://doi.org/10.1016/j.joep.2020.102293.

Engel, C., & Rockenbach, B. (2009). We are not alone: The impact of externalities on public good provision. *Preprints of the Max Planck Institute for Research on Collective Goods*, No. 2009/29. https://dx.doi.org/10.2139/ssrn.1463259.

Engel, C., & Rockenbach, B. (2014). Give everybody a voice! The power of voting in a public goods experiment with externalities. *Preprints of the Max Planck Institute for Research on Collective Goods*, No. 2014/16. https://dx.doi.org/10.2139/ssrn.2519479.

Engel, C., & Rockenbach, B. (2024). What makes cooperation precarious?. *Journal of Economic Psychology, 101*, 102712. https://doi.org/10.1016/j.joep.2024.102712.

Engel, C., & Van Lange, P. A. (2021). Social mindfulness is normative when costs are low, but rapidly declines with increases in costs. *Judgment & Decision Making, 16*(2), 290–322. https://doi.org/10.1017/S1930297500008585.

Engel, C., & Zhurakhovska, L. (2011). Oligopoly as a socially embedded dilemma. An experiment. *Preprints of the Max Planck Institute for Research on Collective Goods*, No. 2011/1. https://dx.doi.org/10.2139/ssrn.1750094.

Engel, C., & Zhurakhovska, L. (2014). Conditional cooperation with negative externalities – An experiment. *Journal of Economic Behavior and Organization, 108*, 252–260. https://doi.org/10.1016/j.jebo.2014.10.005.

Engel, C., & Zhurakhovska, L. (2016). When is the risk of cooperation worth taking? The prisoner's dilemma as a game of multiple motives. *Applied Economics Letters, 23*(16), 1157–1161. https://doi.org/10.1080/13504851.2016.1139672.

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, *90*(4), 980–994. https://doi.org/10.1257/aer.90.4.980.

Fischbacher, U., & Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review*, *100*(1), 541–556. https://doi.org/10.1257/aer.100.1.541.

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, *71*(3), 397–404. https://doi.org/10.1016/S0165-1765(01)00394-9.

Friedman, D., & Oprea, R. (2012). A continuous dilemma. *American Economic Review*, *102*(1), 337–363. https://doi.org/10.1257/aer.102.1.337.

Hertwig, R., & Engel, C. (2016). Homo ignorans: Deliberately choosing not to know. *Perspectives on Psychological Science*, *11*(3), 359–372. https://doi.org/10.1177/1745691616635594.

Hertwig, R., & Engel, C. (2021). Homo ignorans: Deliberately choosing not to know. In R. Hertwig & C. Engel (Eds.), *Deliberate ignorance: Choosing not to know* (pp. 3–18). Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/13757.003.0004.

Hertwig, R., & Grüne-Yanoff, T. (2017). Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science, 12*(6), 973–986. https://doi.org/10.1177/1745691617702496.

Offer, K., Mischkowski, D., Rahwan, Z., & Engel, C. (2024). Deliberately ignoring inequality to avoid rejecting unfair offers. *Communications Psychology, 2*(48), 1–13. https://doi.org/10.1038/s44271-024-00093-6.

Teichman, D., Talley, E., Egidy, S., Engel, C., Gummadi, K. P., Hagel, K., Lewandowsky, S., MacCoun, R. J., Utz, S., & Zamir, E. (2021). Institutions promoting or countering deliberate ignorance. In R. Hertwig & C. Engel (Eds.), *Deliberate ignorance: Choosing not to know* (pp. 275–298). Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/13757.003.0004.

Van Doesum, N. J., Van Lange, P. A., Mischkowski, D., & Van Dijk, E. (2025). Social mindfulness and the SoMi Paradigm: A decade of research on low-cost prosociality. *European Review of Social Psychology*, 1–51. https://doi.org/10.1080/10463283.2025.2457918.