

Intuition, deliberation, and the evolution of cooperation

Adam Bear^{a,1} and David G. Rand^{a,b,c,1}

^aDepartment of Psychology, Yale University, New Haven, CT 06511; ^bDepartment of Economics, Yale University, New Haven, CT 06511; and ^cSchool of Management, Yale University, New Haven, CT 06511

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved November 24, 2015 (received for review September 5, 2015)

Humans often cooperate with strangers, despite the costs involved. A long tradition of theoretical modeling has sought ultimate evolutionary explanations for this seemingly altruistic behavior. More recently, an entirely separate body of experimental work has begun to investigate cooperation's proximate cognitive underpinnings using a dual-process framework: Is deliberative self-control necessary to reign in selfish impulses, or does self-interested deliberation restrain an intuitive desire to cooperate? Integrating these ultimate and proximate approaches, we introduce dual-process cognition into a formal game-theoretic model of the evolution of cooperation. Agents play prisoner's dilemma games, some of which are one-shot and others of which involve reciprocity. They can either respond by using a generalized intuition, which is not sensitive to whether the game is one-shot or reciprocal, or pay a (stochastically varying) cost to deliberate and tailor their strategy to the type of game they are facing. We find that, depending on the level of reciprocity and assortment, selection favors one of two strategies: intuitive defectors who never deliberate, or dual-process agents who intuitively cooperate but sometimes use deliberation to defect in one-shot games. Critically, selection never favors agents who use deliberation to override selfish impulses: Deliberation only serves to undermine cooperation with strangers. Thus, by introducing a formal theoretical framework for exploring cooperation through a dual-process lens, we provide a clear answer regarding the role of deliberation in cooperation based on evolutionary modeling, help to organize a growing body of sometimes-conflicting empirical results, and shed light on the nature of human cognition and social decision making.

dual process | cooperation | evolutionary game theory | prisoner's dilemma | heuristics

Cooperation, where people pay costs to benefit others, is a defining feature of human social interaction. However, our willingness to cooperate is puzzling because of the individual costs that cooperation entails. Explaining how the “selfish” process of evolution could have given rise to seemingly altruistic cooperation has been a major focus of research across the natural and social sciences for decades. Using the tools of evolutionary game theory, great progress has been made in identifying mechanisms by which selection can favor cooperative strategies, providing ultimate explanations for the widespread cooperation observed in human societies (1).

In recent years, the proximate cognitive mechanisms underpinning human cooperation have also begun to receive widespread attention. For example, a wide range of experimental evidence suggests that emotion and intuition play a key role in motivating cooperation (2–5). The dual-process perspective on decision making (6–8) offers a powerful framework for integrating these observations. In the dual-process framework, decisions are conceptualized as arising from competition between two types of cognitive processes: (i) automatic, intuitive processes that are relatively effortless but inflexible; and (ii) controlled, deliberative processes that are relatively effortful but flexible. In many situations, intuitive and deliberative processes can favor different decisions, leading to inner conflict: Rather than being of a single mind, people are torn between competing desires.

Despite the widespread attention that dual-process theories have received in the psychological and economic sciences (including incorporation into formal decision making models; refs. 9–11); the existence of related discussion in the theoretical biology literature regarding error management (12–14), tradeoffs between fixed and flexible behaviors (15–18), and cultural evolution and norm internalization (2, 19, 20); and a long interdisciplinary tradition of arguments suggesting that strategies developed in repeated interactions spill over to influence behavior in one-shot anonymous settings (21–25), the dual-process framework has been almost entirely absent from formal models of the evolution of cooperation. Traditional evolutionary game theory models of cooperation focus on behavior, rather than the cognition that underlies behavior. Therefore, these models do not shed light on when selection may favor the use of intuition versus deliberation, or which specific intuitive and deliberative responses will be favored by selection.

In this paper, we build a bridge between ultimate and proximate levels of analysis to address these questions, introducing an evolutionary game-theoretic model of cooperation that allows for dual-process agents. These agents interact in a varied social environment, where interactions differ in the extent to which current actions carry future consequences. To capture the tradeoff between flexibility and effort that is central to many dual-process theories, we allow our agents to either (i) use an intuitive response that is not sensitive to the type of interaction currently faced; or (ii) pay a cost to deliberate, tailoring their action to the details of the current interaction.

Significance

The role of intuition versus deliberation in human cooperation has received widespread attention from experimentalists across the behavioral sciences in recent years. Yet a formal theoretical framework for addressing this question has been absent. Here, we introduce an evolutionary game-theoretic model of dual-process agents playing prisoner's dilemma games. We find that, across many types of environments, evolution only ever favors agents who (i) always intuitively defect, or (ii) are intuitively predisposed to cooperate but who, when deliberating, switch to defection if it is in their self-interest to do so. Our model offers a clear explanation for why we should expect deliberation to promote selfishness rather than cooperation and unifies apparently contradictory empirical results regarding intuition and cooperation.

Author contributions: A.B. and D.G.R. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: Code implementing the model in MATLAB is available at <https://gist.github.com/adambear91/c9b3c02a7b9240e288cc>.

¹To whom correspondence may be addressed. Email: adam.bear@yale.edu or david.rand@yale.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1517780113/-DCSupplemental.

We then use this framework to explore the consequences of reciprocity and assortment (26, 27), two of the most widely studied mechanisms for the evolution of cooperation. We ask when (and to what extent) agents evolve to pay the cost of deliberation; when evolution favors intuitive responses that are selfish versus cooperative; and whether deliberation serves to increase or decrease social welfare. In doing so, we provide a formal theoretical framework to guide the emerging body of empirical work exploring prosociality from a dual-process perspective, and provide insight into the cognitive underpinnings of human cooperation.

Model

There are two key dimensions on which our model differs from typical models of the evolution of cooperation: (i) in each generation, agents play more than one type of game; and (ii) agents need not have a single fixed strategy, but can engage in costly deliberation to tailor their response to the type of game they are facing.

With respect to multiple game types, our agents face both one-shot anonymous prisoner's dilemma (PD) games (which occur with probability $1-p$) and PDs where reciprocal consequences exist (which occur with probability p). In the one-shot PDs, agents can cooperate by paying a cost c to give a benefit b to their partner, or defect by doing nothing. In the games with reciprocal consequences, we capture the core of reciprocity [be it via repeated interactions, reputation effects, or sanctions (1)] by modifying the off-diagonal elements of the PD payoff structure: When exploitation occurs, such that one player defects while the other cooperates, the benefit to the defector is reduced (due to, e.g., lost future cooperation, damaged reputation, or material punishment), as is the cost to the cooperator (due to, e.g., switching to defection, improved reputation, or material rewards). As a result, the social dilemma of the PD is transformed into a coordination game: It becomes payoff-maximizing to cooperate if one's partner also cooperates. For simplicity, we focus on the limiting case where when one player cooperates and the other defects, both receive zero payoffs. Because this simplified payoff structure is analogous to the average payoff per round of an infinitely repeated PD between "tit-for-tat" and "always defect," for expositional purposes, we refer to games with reciprocal consequences as "repeated games." Critically, however, our results do not rely on this simplifying assumption, or on the use of repeated games more generally (mitigating potential concerns about alternative repeated game strategy sets; ref. 28). Rather, they hold whenever agents face any of a broad class of cooperative coordination games with probability p ; see *SI Appendix, Section 6* for details. Similarly, the social dilemma that occurs with probability $1-p$ need not be a one-shot PD—equivalent results would be obtained by using any game where cooperation always earns less than noncooperation.

We also consider the other main force that has been argued to underlie the evolution of human cooperation: assortment (29). An agent plays against another agent having the same strategy as herself with probability a and plays with an agent selected at random from the population with probability $1-a$. Thus, a captures the extent to which agents of similar types are more likely than chance to interact. This assortment could arise from relatedness, spatial or networked interactions, or group selection (1).

With respect to multiple strategies within a single agent, our model allows agents to use two forms of decision making: intuition or deliberation (see Fig. 1 for a visual depiction of an agent's decision process; and *SI Appendix, Section 1* for further details). Among the various dimensions upon which these modes of cognitive processing differ (6), we focus on the fact that intuitive responses are quick and relatively effortless (and thus less costly), but also less sensitive to situational and strategic details than deliberative responses. For simplicity, we focus on the limiting case where intuition is totally inflexible and deliberation is perfectly flexible/accurate. When agents decide intuitively, they cooperate with some fixed probability S_i , regardless of whether the game is

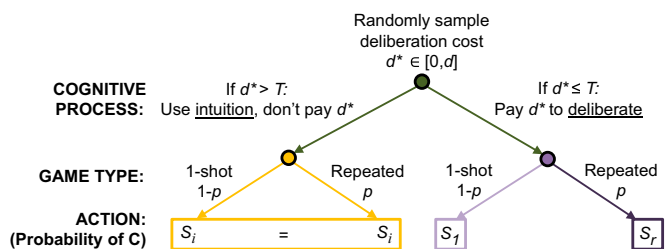


Fig. 1. Agents play PD games that are either one-shot or involve reciprocity, and either use a generalized intuitive strategy that does not depend on game type, or engage in costly deliberation and tailor their strategy based on game type. The strategy space for the agents in our model, which consists of four variables T , S_i , S_r , and S_c , is visualized here along with the sequence of events within each interaction between two agents (both agents face the same decision, so for illustrative simplicity only one agent's decision process is shown). First, the agent's cost of deliberation for this interaction d^* is sampled uniformly from the interval $[0, d]$. The agent's deliberation threshold T then determines which mode of cognitive processing is applied. If $d^* > T$, it is too costly to deliberate in this interaction and she makes her cooperation decision based on her generalized intuitive response S_i ; intuition cannot differentiate between game types, and so regardless of whether the game is one-shot (probability $1-p$) or repeated (probability p), she plays the cooperative strategy with probability S_i . If $d^* \leq T$, however, deliberation is not too costly, so she pays the cost d^* and uses deliberation to tailor her play to the type of game she is facing: If the game is one-shot, she plays the cooperative strategy with probability S_i , and if the game is repeated, she plays the cooperative strategy with probability S_r . For example, when deliberating, an agent could decide to defect in a one-shot game ($S_i = 0$) but cooperate in a repeated game ($S_r = 1$). In contrast, when using intuition, this agent must either cooperate in both contexts ($S_i = 1$) or defect in both contexts ($S_i = 0$).

one-shot or repeated. Deliberating, conversely, allows agents to potentially override this intuitive response and tailor their strategy to the type of game they are facing. When deliberating, agents cooperate with probability S_i if the game they are facing is one-shot, and cooperate with probability S_r if it is repeated.

The flexibility of deliberation, however, comes at a cost (30, 31). This cost can take several forms. First, deliberation is typically slower than intuition, and this greater time investment can be costly. For example, sometimes decisions must be made quickly lest you miss out on the opportunity to act. Second, deliberation is more cognitively demanding than intuition: Reasoning your way to an optimal solution takes cognitive effort. Furthermore, because intuitions are typically low-level cognitive processes that are triggered automatically and reflexively (7, 8), cognitive resources may also be required to inhibit intuitive responses when deliberation reveals that they are suboptimal. These cognitive demands associated with deliberation can impose fitness costs by reducing the agent's ability to devote cognitive resources to other important tasks unrelated to the cooperation decision. The fitness costs associated with this need to redirect cognitive resources are particularly large when agents are under cognitive load or are fatigued.

Thus, deliberation is costly, but the size of that cost varies from decision to decision (between 0 and some maximum value d). For simplicity, in each interaction, we independently sample a cost of deliberation d^* for each agent from a uniform distribution over $[0, d]$.

In addition to evolving different intuitive and deliberative responses, we allow natural selection to act on the extent to which agents rely on intuition versus deliberation. Specifically, each agent's strategy specifies a deliberation cost threshold T , such that they deliberate in interactions where the deliberation cost is sufficiently small, $d^* \leq T$, but act intuitively when deliberation is sufficiently costly, $d^* > T$. Thus, in any given interaction, an agent with threshold T deliberates with probability T/d and uses intuition with probability $1-T/d$. The higher an agent's value of T , the more that agent tends to deliberate.

one-shot, $S_I = 0$. In other words, DC players use deliberation to override their cooperative intuitions when they find themselves in a one-shot game, and instead defect.

Together, these two Nash equilibria reproduce the pattern observed in the evolutionary dynamics (Fig. 2C). The transition from ID to DC occurs precisely at the point where DC risk-dominates ID (i.e., where DC earns more than ID in a population where both strategies are equally likely, which is known to predict evolutionary success; see *SI Appendix, Section 3* for details), after which point mean $T = c(1-p)$ (declining linearly in p). Furthermore, after this point, increasing the probability of games being repeated p has no effect on cooperation in repeated games (which is at ceiling), but instead increases cooperation in one-shot games (Fig. 2D): Across a wide range of parameters, cooperation in repeated games is high, and cooperation in one-shot games is lower but substantially greater than zero [as is typical of human behavior (1)].

Why is $T = c(1-p)$ the maximum cost for which DC will deliberate? Deliberation allows DC to avoid cooperating (and, thus, avoid incurring a cost c) in the fraction $(1-p)$ of interactions that are one-shot; in the fraction p of interactions that are repeated, there is no benefit to deliberating, because DC's intuitive and deliberative responses agree on cooperating. Therefore, $c(1-p)$ is DC's expected payoff gain from deliberating, and so deliberation is disadvantageous when it is more costly than $c(1-p)$. This condition emphasizes the fact that deliberation's only function for DC agents is to restrain the impulse to cooperate in one-shot games. Intuition, conversely, functions as a "repeated game" social heuristic (24, 25), prescribing the cooperative strategy that is typically advantageous (given the sufficiently high prevalence of repeated games).

Critically, we do not observe an equilibrium that intuitively defects but uses deliberation to cooperate in repeated games. This strategy is not an equilibrium because of a coordination problem: It is only beneficial to override a defecting intuition to cooperate in a repeated game when your partner also plays a cooperative strategy. Thus, ID's expected payoff from deliberating when playing a repeated game with an ID partner is discounted by the extent to which their partner fails to deliberate (and so the partner defects, even though the game is repeated). As a result, IDs maximize their payoff by deliberating less than their partners, leading to an unraveling of deliberation: Any nonzero deliberation threshold is not Nash, because there is an incentive to deviate by deliberating less than your partner. (This is not true for the intuitively cooperative strategy DC deliberating in one-shot games, because in one-shot games it is always beneficial to switch to defection, no matter what the other agent does.)

Finally, we investigate the effect of assortment in our model. Nash equilibrium analysis again shows that ID and DC are the only equilibria, with DC's deliberation threshold now being $T = (c-ba)(1-p)$. Increasing a acts in a similar way to increasing p , allowing DC to be favored over ID and, subsequently, reducing T (Fig. 3A). Consistent with this analysis, evolutionary calculations show an interaction between a and p . When repeated games are rare (small p), increasing a allows intuitive cooperation to succeed and initially increases T (as DC begins to outperform ID); as a increases further, however, T decreases (Fig. 3B). When repeated games are common (large p), conversely, DC is dominant even without assortment; therefore, increasing a always decreases T (Fig. 3C). These analyses show that our results are robust to assumptions about the evolutionary history of humans: Regardless of whether most interactions involved reciprocity with little assortment, or most interactions were one-shot but assorted, selection favors the same intuitively cooperative dual-process strategy (and never a strategy that uses deliberation to cooperate by overruling intuitive defection).

Discussion

By integrating dual-process cognition into a game-theoretic model of the evolution of cooperation based on reciprocity and assortment, we provide a formal theoretical framework for considering the

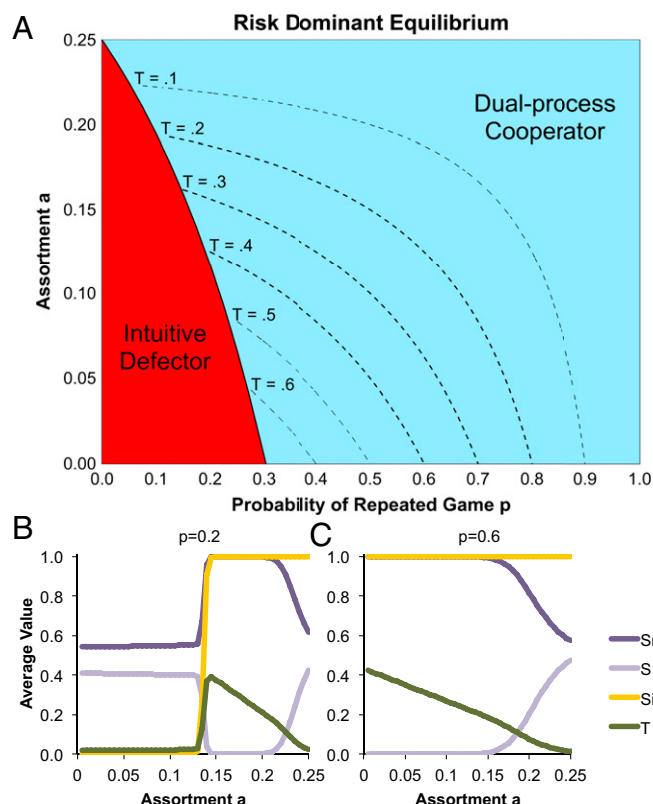


Fig. 3. Assortment also favors the evolution of dual process cooperators. (A) Nash equilibrium calculations with assortment ($a > 0$) again find that ID and DC are the only possible equilibria. The risk-dominant strategy is shown as a function of the probability of repeated games p and the level of assortment a (risk-dominance indicates which strategy will be favored by selection; see *SI Appendix, Fig. S1* for corresponding evolutionary calculations). DC is favored so long as either p or a are sufficiently large. Thus, regardless of whether evolution occurs in a world where reciprocity is powerful and assortment is weak, or where reciprocity is weak and assortment is strong, selection favors intuitive cooperation combined with deliberative defection in one-shot games. Also shown are isoclines for T within the region where DC is favored. Here, increasing either a or p decreases T . Thus, we find an interaction with reciprocity when considering how assortment affects cognitive style: when p is low, assortment initially increases deliberation, but when p is high, assortment monotonically decreases deliberation. This interaction is visualized by showing the average values of each strategy variable in the steady-state distribution of the evolutionary process as a function of assortment a , for $p = 0.2$ (B) and $p = 0.6$ (C). Evolutionary calculations use $N = 50$, $b = 4$, $c = 1$, $d = 1$, and $w = 6$.

question of whether prosociality is intuitive or whether it requires self-control. We find that evolution never favors strategies for which deliberation increases cooperation. Instead, when deliberation occurs, it always works to undermine cooperation in one-shot interactions. Intuition, conversely, acts as a social heuristic (24, 25), implementing the behavior that is typically advantageous (cooperation, unless p and a are both sufficiently small, because the cost of missing out on reciprocal cooperation outweighs the cost of needlessly cooperating in one-shot games (14). Thus, our model helps to explain why people cooperate even in one-shot anonymous settings, but less frequently than they do in repeated interactions. Furthermore, we offer an explanation for why cooperation in such situations is typically "conditional" rather than "unconditional" (32) (i.e., why people will cooperate in one-shot games, but only if they expect their partner to also cooperate): when one-shot cooperation evolves in our model, it occurs via an intuitive response that treats social dilemmas as if they were coordination games.

Our model also makes numerous clear, testable predictions about human cognition. First, in one-shot anonymous interactions,

promoting intuition (increasing the cost of deliberation d^*) should, on average, increase cooperation relative to promoting deliberation (reducing d^*). This prediction holds even in laboratory experiments where participants are explicitly told that the game they are playing is one-shot, for at least two reasons. First, deliberation is required to translate this explicit knowledge about game length into a strategic understanding that cooperative intuitions should be overridden (many participants mistakenly believe that it is in their self-interest to cooperate even when they are told the game is one-shot, as shown for example in refs. 33 and 34). Second, further cognitive effort may be required to inhibit the intuitive response because it is triggered automatically. In line with this prediction, data from numerous research groups show that experimentally inducing participants to decide more intuitively using time pressure (4, 24, 33), cognitive load (35–37), conceptual inductions (3, 4, 38), or variation in payment delays (11, 39) can increase prosociality in one-shot economic games.

Furthermore, our model predicts substantial heterogeneity across individuals in this effect. People who developed their strategies in social settings with little future consequence for bad behavior (small p) and low levels of assortment (small a) are predicted to intuitively defect ($S_i = 0$), and to engage in little deliberation, regardless of the cost ($T = 0$). Thus, experimentally manipulating the cost of deliberation should not affect these participants' cooperation. Consistent with this prediction, time constraint manipulations were found to have little effect on participants with untrustworthy daily-life interaction partners (34) or participants from a country with low levels of interpersonal trust and cooperation (40). Furthermore, our model predicts that when p and/or a becomes sufficiently large, T also approaches 0 (albeit with a cooperative intuition, $S_i = 1$). Therefore, people who developed their strategies in contexts that were extremely favorable to cooperation should also be relatively unaffected by cognitive process manipulations. This prediction may help to explain why some one-shot game studies find no effect of manipulating intuition (41). Further support for the predicted link between extent of future consequences and intuitive one-shot cooperation comes from laboratory evidence for “habits of virtue,” where repeated game play spills over into subsequent one-shot interactions, but only among participants who rely on heuristics (25). These various experience-related results emphasize that our model operates in the domain of cultural evolution and social learning (1, 19, 20), in addition to (or instead of) genetic evolution.

Our model also predicts variation in the effect of intuition versus deliberation across contexts: Whereas deliberating undermines cooperation in one-shot games, it is predicted to have no effect in repeated games. The DC strategy's cooperative intuition is supported (and, therefore, unaffected) by deliberation in repeated games; and the ID strategy defects under all circumstances, and so is unaffected by deliberation. Consistent with this prediction, manipulating intuition had little effect on cooperation in a repeated four-player PD (42) or a modified public goods game where cooperating was individually payoff-maximizing (34).

Although our model predicts that manipulating the use of intuition versus deliberation will have the aforementioned effects, it conversely predicts that there will likely not be a consistent correlation between one-shot cooperation and an individual's willingness to deliberate T (i.e., their “cognitive style”): Highly intuitive (low T) individuals can either be intuitive defectors or intuitive cooperators. In line with this prediction, little consistent association has been found between an individual's cognitive style and their overall willingness to cooperate in one-shot games (4, 43, 44). Furthermore, individual differences in reaction times, which are often interpreted as a proxy for intuitiveness (although see refs. 45 and 46 for an alternative interpretation based on decision conflict), have been associated with both increased (4, 47–49) and decreased (50, 51) cooperation. Our model therefore helps to explain the otherwise-puzzling difference in experimental results between cognitive process manipulations and reaction time

correlations. Our model also makes the further prediction, untested as far as we know, that a consistent correlation between cognitive style and cooperation should emerge in samples that are restricted to individuals who developed their strategies under conditions where p and/or r were sufficiently large.

The model we have presented here is, in the game-theoretic tradition, highly stylized and focused on limiting cases for tractability. In particular, we assume (i) that agents engage in only two different types of games, rather than, for example, sampling PD game lengths (or coordination game payoffs) from a distribution; (ii) that deliberation is perfectly accurate in assessing the type of game being played, whereas intuition is totally insensitive to game type; and (iii) that the cost of deliberation is sampled from a uniform distribution on the interval $[0, d]$, rather than a more realistic distribution of costs. Future work should extend the framework we introduce here to explore the effects of relaxing these assumptions and incorporate other nuances of dual-process cognition that were not included in this first model. For example, intuitive thinking might be made observable, such that agents could condition on the cognitive style of their partners (as in recent work on “cooperation without looking”; ref. 52). Or, feedback between the population state and the environment (i.e., the model parameters) could be incorporated, as has been done in recent models of the evolution of dual-process agents in the context of intertemporal choice (17, 18).

Future work should also use the framework introduced here to explore the evolution of cognition in domains beyond cooperation. For example, our framework could easily be extended to describe the internalization of a variety of social norms unrelated to cooperation, such as rituals or taboos. Consider any situation in which following the norm is individually costly, but becomes payoff-maximizing when interacting with others who also follow the norm (e.g., because they would sanction norm violations). Our model's logic suggests that selection can favor a strategy that (i) intuitively follows the norm, but (ii) uses deliberation to violate the norm in settings where there is little threat of sanctions (e.g., because one's behavior is unobservable), so long as the overall probability of being sanctioned and/or the severity of the sanctions are sufficiently high.

Our framework could also be extended to explain rejections in the ultimatum game (UG). In this game, Player 1 proposes how a monetary stake should be divided between herself and Player 2. Player 2 can then either accept, or reject such that both players receive nothing. Behavioral experiments suggest that in the one-shot anonymous UG, intuition supports rejecting unfair offers, whereas deliberation leads to increased acceptance (53–55) (although neurobiological evidence is somewhat more mixed; refs. 56 and 57). Such a pattern of intuitive rejection is easily explained by our framework, because rejecting unfair offers is advantageous in repeated games, but costly in one-shot games. Thus, the same logic that leads to selection favoring intuitive cooperation and deliberative defection in our model's one-shot PDs can lead selection to favor intuitive rejection and deliberative acceptance in one-shot UGs.

In sum, we have integrated dual-process theories of cognition from the behavioral sciences with formal game-theoretic models of the evolution of cooperation. Our model shows how it can be adaptive for humans to think both “fast and slow” and provides an explanation for why people sometimes (but not always) cooperate in one-shot anonymous interactions. In doing so, we provide a formal demonstration of how spillovers from settings where cooperation can be payoff-maximizing (e.g., repeated interactions) lead to cooperation in social dilemmas where cooperation is never in one's self-interest (21–25, 58). Although many have suggested that it takes cold, deliberative reasoning to get people to engage in this kind of prosocial behavior, our evolutionary model finds precisely the opposite. It is not reflective thought that allows people to forego their selfish impulses, but rather reflective thought that undermines the impulse to cooperate.

ACKNOWLEDGMENTS. We thank Sam Bowles, Rob Boyd, Molly Crockett, Fiery Cushman, Andy Delton, Josh Greene, Moshe Hoffman, Jillian Jordan, Max

Krasnow, Adam Morris, Martin Nowak, and Erez Yoeli for helpful comments and discussion. Funding was provided by the John Templeton Foundation.

- Rand DG, Nowak MA (2013) Human cooperation. *Trends Cogn Sci* 17(8):413–425.
- Bowles S, Gintis H (2002) Prosocial emotions. *The Economy as an Evolving Complex System 3*, eds Blume LE, Durlauf SN (Oxford Univ Press, Oxford), pp 339–364.
- Small DA, Loewenstein G, Slovic P (2007) Sympathy and callousness: The impact of deliberative thought on donations to identifiable and statistical victims. *Organ Behav Hum Decis Process* 102(2):143–153.
- Rand DG, Greene JD, Nowak MA (2012) Spontaneous giving and calculated greed. *Nature* 489(7416):427–430.
- Zaki J, Mitchell JP (2013) Intuitive Prosociality. *Curr Dir Psychol Sci* 22(6):466–470.
- Kahneman D (2003) A perspective on judgment and choice: Mapping bounded rationality. *Am Psychol* 58(9):697–720.
- Evans JSBT (2008) Dual-processing accounts of reasoning, judgment, and social cognition. *Annu Rev Psychol* 59:255–278.
- Slooman SA (1996) The empirical case for two systems of reasoning. *Psychol Bull* 119(1):3–22.
- Fudenberg D, Levine DK (2006) A dual-self model of impulse control. *Am Econ Rev* 96(5):1449–1476.
- Loewenstein GF, O'Donoghue T (2004) Animal spirits: Affective and deliberative processes in economic behavior. Available at papers.ssrn.com/sol3/papers.cfm?abstract_id=539843. Accessed October 1, 2015.
- Dreber A, Fudenberg D, Levine DK, Rand DG (2015) Self-control, social preferences and the effect of delayed payments. Available at papers.ssrn.com/sol3/papers.cfm?abstract_id=2477454. Accessed October 1, 2015.
- Johnson DDP, Blumstein DT, Fowler JH, Haselton MG (2013) The evolution of error: Error management, cognitive constraints, and adaptive decision-making biases. *Trends Ecol Evol* 28(8):474–481.
- Haselton MG, Nettle D (2006) The paranoid optimist: An integrative evolutionary model of cognitive biases. *Pers Soc Psychol Rev* 10(1):47–66.
- Delton AW, Krasnow MM, Cosmides L, Tooby J (2011) Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proc Natl Acad Sci USA* 108(32):13335–13340.
- Wolf M, Doorn GSV, Weissing FJ (2008) Evolutionary emergence of responsive and unresponsive personalities. *Proc Natl Acad Sci USA*.
- McElreath R, Strimling P (2006) How noisy information and individual asymmetries can make 'personality' an adaptation: A simple model. *Anim Behav* 72(5):1135–1139.
- Tomlin D, Rand DG, Ludvig EA, Cohen JD (2015) The evolution and devolution of cognitive control: The costs of deliberation in a competitive world. *Sci Rep* 5:11002.
- Toupo DFP, Strogatz SH, Cohen JD, Rand DG (2015) Evolutionary game dynamics of controlled and automatic decision-making. *Chaos* 25(7):073120.
- Chudek M, Henrich J (2011) Culture-gene coevolution, norm-psychology and the emergence of human prosociality. *Trends Cogn Sci* 15(5):218–226.
- Boyd R, Richerson PJ (2009) Culture and the evolution of human cooperation. *Philos Trans R Soc Lond B Biol Sci* 364(1533):3281–3288.
- Tooby J, Cosmides L (1990) The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethol Sociobiol* 11(4):375–424.
- Hagen EH, Hammerstein P (2006) Game theory and human evolution: A critique of some recent interpretations of experimental games. *Theor Popul Biol* 69(3):339–348.
- Yamagishi T (2007) The social exchange heuristic: A psychological mechanism that makes a system of generalized exchange self-sustaining. *Cultural and Ecological Foundation of the Mind*, eds Radford M, Ohnuma S, Yamagishi T (Hokkaido Univ Press, Sapporo, Japan), pp 11–37.
- Rand DG, et al. (2014) Social heuristics shape intuitive cooperation. *Nat Commun* 5:3677.
- Peysakhovich A, Rand DG (September 9, 2015) Habits of virtue: Creating norms of cooperation and defection in the laboratory. *Manage Sci*, 10.1287/mnsc.2015.2168.
- Axelrod R, Hamilton WD (1981) The evolution of cooperation. *Science* 211(4489):1390–1396.
- van Veelen M, García J, Rand DG, Nowak MA (2012) Direct reciprocity in structured populations. *Proc Natl Acad Sci USA* 109(25):9929–9934.
- McNally L, Tanner CJ (2011) Flexible strategies, forgiveness, and the evolution of generosity in one-shot encounters. *Proc Natl Acad Sci USA* 108(44):E971, author reply E972.
- Fletcher JA, Doebeli M (2009) A simple and general explanation for the evolution of altruism. *Proc Biol Sci* 276(1654):13–19.
- Posner MI, Snyder CRR (1975) Facilitation and inhibition in the processing of signals. *Attention and Performance*, ed Rabbit PMA (Academic, London) Vol 5, 669–682.
- Kool W, McGuire JT, Rosen ZB, Botvinick MM (2010) Decision making and the avoidance of cognitive demand. *J Exp Psychol Gen* 139(4):665–682.
- Fischbacher U, Gächter S, Fehr E (2001) Are people conditionally cooperative? Evidence from a public goods experiment. *Econ Lett* 71(3):397–404.
- Rand DG, Newman GE, Wurzbacher O (2015) Social context and the dynamics of cooperative choice. *J Behav Decis Making* 28(2):159–166.
- Rand DG, Kraft-Todd GT (2014) Reflection does not undermine self-interested prosociality. *Front Behav Neurosci* 8:300.
- Schulz JF, Fischbacher U, Thöni C, Utikal V (2014) Affect and fairness: Dictator games under cognitive load. *J Econ Psychol* 41:77–87.
- Cornelissen G, Dewitte S, Warlop L (2011) Are social value orientations expressed automatically? Decision making in the dictator game. *Pers Soc Psychol Bull* 37(8):1080–1090.
- Roch SG, Lane JAS, Samuelson CD, Allison ST, Dent JL (2000) Cognitive load and the equality heuristic: A two-stage model of resource overconsumption in small groups. *Organ Behav Hum Decis Process* 83(2):185–212.
- Lotz S (2015) Spontaneous giving under structural inequality: Intuition promotes cooperation in asymmetric social dilemmas. *PLoS One* 10(7):e0131562.
- Kovarik J (2009) Giving it now or later: Altruism and discounting. *Econ Lett* 102(3):152–154.
- Capraro V, Coccioni G (2015) Social setting, intuition and experience in laboratory experiments interact to shape cooperative decision-making. *Proc Biol Sci* 282(11):20150237.
- Tinghög G, et al. (2013) Intuition and cooperation reconsidered. *Nature* 498(7452):E1–E2, discussion E2–E3.
- Duffy S, Smith J (2014) Cognitive load in the multi-player prisoner's dilemma game: Are there brains in games? *J Behav Exper Econ* 51:47–56.
- Arruñada B, Casari M, Pancotto F (2015) Pro-sociality and strategic reasoning in economic decisions. *Front Behav Neurosci* 9:140.
- Yamagishi T, Li Y, Takagishi H, Matsumoto Y, Kiyonari T (2014) In search of Homo economicus. *Psychol Sci* 25(9):1699–1711.
- Evans AM, Dillon KD, Rand DG (2015) Fast but not intuitive, slow but not reflective: Decision conflict drives reaction times in social dilemmas. *J Exp Psychol Gen* 144(5):951–966.
- Krajibich I, Bartling B, Hare T, Fehr E (2015) Rethinking fast and slow based on a critique of reaction-time reverse inference. *Nat Commun* 6:7455.
- Lotito G, Migheili M, Ortona G (2013) Is cooperation instinctive? Evidence from the response times in a public goods game. *J Bioeconomics* 15(2):123–133.
- Cappelen AW, Nielsen UH, Tungodden B, Tyrann JR, Wengström E (September 3, 2015) Fairness is intuitive. *Exp Econ*, 10.1007/s10683-015-9463-y.
- Nielsen UH, Tyrann JR, Wengström E (2014) Second thoughts on free riding. *Econ Lett* 122(2):136–139.
- Piovesan M, Wengström E (2009) Fast or fair? A study of response times. *Econ Lett* 105(2):193–196.
- Fiedler S, Glöckner A, Nicklisch A, Dickert S (2013) Social Value Orientation and information search in social dilemmas: An eye-tracking analysis. *Organ Behav Hum Decis Process* 120(2):272–284.
- Hoffman M, Yoeli E, Nowak MA (2015) Cooperate without looking: Why we care what people think and not just what they do. *Proc Natl Acad Sci USA* 112(6):1727–1732.
- Grimm V, Mengel F (2011) Let me sleep on it: Delay reduces rejection rates in ultimatum games. *Econ Lett* 111(2):113–115.
- Sutter M, Kocher M, Strauß S (2003) Bargaining under time pressure in an experimental ultimatum game. *Econ Lett* 81(3):341–347.
- Halali E, Bereby-Meyer Y, Meiran N (2014) Between self-interest and reciprocity: The social bright side of self-control failure. *J Exp Psychol Gen* 143(2):745–754.
- Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD (2003) The neural basis of economic decision-making in the Ultimatum Game. *Science* 300(5626):1755–1758.
- Knoch D, Pascual-Leone A, Meyer K, Treyer V, Fehr E (2006) Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314(5800):829–832.
- Tomasello M, Melis AP, Tennie C, Wyman E, Herrmann E (2012) Two key steps in the evolution of human cooperation. *Curr Anthropol* 53(6):673–692.

Supplementary Information for Intuition, deliberation, and the evolution of cooperation

Adam Bear, David G. Rand

Contents

1	Strategy space and payoff function	2
2	Nash equilibrium calculations	3
2.1	Setup	3
2.2	Intuitively defecting equilibria	4
2.3	Purely deliberative equilibrium	5
2.4	Intuitively cooperating equilibria	5
2.5	Summary of Nash results	6
2.6	Why is there no equilibrium with $S_i = 0$ and $T > 0$?	7
2.7	Nash calculations with assortment	7
3	Risk dominance calculations	8
3.1	Without assortment	8
3.2	With assortment	9
4	Evolutionary dynamics	9
4.1	Basic setup	9
4.2	Limit of low mutation calculation method	10
5	Robustness of evolutionary results	11
5.1	Robustness to parameter variation	11
5.2	Robustness to higher mutation rates and probabilistic strategies	12
6	Generalized coordination game analysis	12
6.1	Setup	12
6.2	Intuitively cooperating equilibria	13
6.3	Purely deliberative equilibrium	13
6.4	Intuitively defecting equilibria	13
6.5	Summary	14
6.6	Application to repeated PD with finite continuation probability	14
6.7	Application to general PD with reciprocal consequences	15

1 Strategy space and payoff function

Our main model considers agents playing 1-shot Prisoner's Dilemma (PD) games or PDs with reciprocal consequences (modeled using the framework of infinitely repeated games); and responding using either a generalized intuition S_i or paying a cost d^* (stochastically sampled from the interval $[0, d]$) to deliberate and tailor their strategy such that they use strategy S_r if the game is repeated and S_1 if the game is 1-shot. In each interaction, agents choose either the cooperative strategy tit-for-tat (TFT) or the non-cooperative strategy always defect (ALLD). Importantly, as we demonstrate below in Section 6, our results are not specific to TFT and ALLD playing repeated PDs, but instead generalize to a wide range of coordination games.

An agent's strategy profile is specified by four variables: their (i) probability of intuitively playing TFT S_i , (ii) probability of playing TFT when they deliberate and face a repeated game S_r , (iii) probability of playing TFT (i.e. cooperating) when they deliberate and face a 1-shot game S_1 , and (iv) maximum acceptable cost of deliberation T . Since we stipulate that the cost of deliberation is sampled uniformly from the interval $[0, d]$, an agent with threshold T deliberates with probability $\frac{T}{d}$ and on average pays a cost $\frac{T}{2}$ when deliberating. We use a uniform distribution for simplicity, but more realistic cost distributions should not change our results qualitatively.

Here we specify the expected payoff $\pi(x, y)$ of an agent with strategy profile $x = [S_i, S_1, S_r, T]$ playing against an agent with strategy profile $y = [S'_i, S'_1, S'_r, T']$. To do so, we calculate agent x 's expected payoff from playing infinitely repeated PDs with probability p and 1-shot PDs with $1 - p$, over the cases in which (i) both agents deliberate (probability $\frac{TT'}{d^2}$), (ii) agent x deliberates and agent y decides intuitively (probability $\frac{T}{d}(1 - \frac{T'}{d})$), (iii) agent x decides intuitively and agent y deliberates (probability $(1 - \frac{T}{d})\frac{T'}{d}$), and (iv) both agents decide intuitively (probability $(1 - \frac{T}{d})(1 - \frac{T'}{d})$):

$$\pi(x, y) = \frac{TT'}{d^2}(\pi_{DD} - \frac{T}{2}) + \frac{T}{d}(1 - \frac{T'}{d})(\pi_{DI} - \frac{T}{2}) + (1 - \frac{T}{d})\frac{T'}{d}\pi_{ID} + (1 - \frac{T}{d})(1 - \frac{T'}{d})\pi_{II}$$

where π_{DD} is agent x 's expected payoff when both agents deliberate, π_{DI} is agent x 's expected payoff when agent x deliberates and agent y uses intuition, and so on.

These expected payoffs are calculated based on the payoff tables for 1-shot and repeated PDs. In 1-shot games, TFT cooperates and pays a cost c to give a benefit b to the partner, while ALLD defects and does nothing. Thus, the payoff table for the 1-shot games is given by

1-shot PD Payoffs		
	TFT	ALLD
TFT	$b - c$	$-c$
ALLD	b	0

where the row player's payoff is shown.

To make payoffs in an infinitely repeated game comparable to those of a 1-shot game, we use the average payoff per round. Here, two TFT agents cooperate with each other in every round and earn average payoffs per round of $b - c$, while two ALLD agents defect every round, earning 0.

Thus these payoffs are the same as the 1-shot PD. When a TFT agent and an ALLD agent meet, however, the outcome differs from the 1-shot game, because the TFT agent cooperates only on the first round, and then defects in every subsequent round. Because the interaction is modeled as being infinitely repeated, the first round (where TFT cooperates) contributes only a negligible amount to the average payoff. Therefore, both agents earn an average payoff per round of 0. Therefore the payoff table for the infinitely repeated PD is given by

Infinitely Repeated PD Payoffs

	TFT	ALLD
TFT	$b - c$	0
ALLD	0	0

where $b, c > 0$.

Importantly, using total payoff (rather than average payoff per round) in a game with a finite continuation probability, such that the first round does influence payoffs and causes some negative cost for TFT and positive benefit for ALLD, does not qualitatively change our results; see Section 6 below.

Substituting in relevant payoff values yields

$$\begin{aligned}
\pi_{DD} &= p(S_r S'_r (b - c)) + (1 - p)(S_1 S'_1 (b - c) + S_1 (1 - S'_1)(-c) + (1 - S_1) S'_1 b) \\
\pi_{DI} &= p(S_r S'_i (b - c)) + (1 - p)(S_1 S'_i (b - c) + S_1 (1 - S'_i)(-c) + (1 - S_1) S'_i b) \\
\pi_{ID} &= p(S_i S'_r (b - c)) + (1 - p)(S_i S'_1 (b - c) + S_i (1 - S'_1)(-c) + (1 - S_i) S'_1 b) \\
\pi_{II} &= p(S_i S'_i (b - c)) + (1 - p)(S_i S'_i (b - c) + S_i (1 - S'_i)(-c) + (1 - S_i) S'_i b)
\end{aligned}$$

2 Nash equilibrium calculations

2.1 Setup

To facilitate Nash equilibria calculations, we consider a strategy space which is simplified relative to the main model in two ways: (i) agents' intuitive response S_i is limited to being either 0 (never play TFT) or 1 (always play TFT); and (ii) agents' deliberative responses are fixed to be $S_1 = 0$ and $S_r = 1$; i.e., always defecting when deliberating and facing a 1-shot game, and always playing TFT when deliberating and facing a repeated game. As in the main model, agents specify a maximum cost of deliberation T ($0 \leq T \leq d$) that they are willing to pay in order to deliberate, and this determines when they deliberate.

Thus, an agent's strategy profile is specified by two variables: 1) a binary variable S_i indicating whether or not the agent intuitively plays the cooperative strategy and 2) a continuous variable T indicating the agent's maximum cost they are willing to pay to deliberate. We denote a strategy profile for this reduced strategy space as $x = [S_i, T]$. (These simplifications of the intuitive and deliberative strategy spaces are justified by our evolutionary simulations using the full strategy space, whose results are in agreement with the results of the Nash calculation for the simplified

strategy space – see main text Figure 2.)

A strategy profile x is a Nash equilibrium if no strategy profile y is able to get a higher payoff against x than x gets against itself. That is,

$$\forall y : \pi(x, x) \geq \pi(y, x).$$

Given our restricted strategy space, the set of possible strategy profiles that an agent can adopt can be thought of as two continuous sets: 1) the set of strategy profiles that intuitively defect ($S_i = 0$) and have threshold $0 \leq T \leq d$, and 2) the set of strategy profiles that intuitively cooperate ($S_i = 1$) and have threshold $0 \leq T \leq d$.

2.2 Intuitively defecting equilibria

We first consider whether any strategy profile with $S_i = 0$ is a Nash. To do this, we calculate the expression for the payoff that an agent with $S_i = 0$ and $T = T$ gets against an agent with $S_i = 0$ and $T = T'$:

$$\pi([0, T], [0, T']) = \frac{T^2(\frac{T'}{d} - 1)}{2d} - \frac{TT'(\frac{T'}{2} - p(b - c))}{d^2}.$$

Since the concavity of this function (with respect to T) is always negative ($\frac{\partial^2}{\partial T^2} \pi([0, T], [0, T']) < 0$), there is a unique best-response $[0, T_b]$ that maximizes one's payoff when playing against $[0, T']$, which can be found by asking what value of T satisfies the equation

$$\frac{\partial}{\partial T} \pi([0, T], [0, T']) = 0.$$

Doing so yields

$$T_b = \frac{p(b - c)T'}{d}.$$

Since a strategy profile must be a best response against itself in order to be Nash, it must be the case that $T_b = T'$ in the above equation for T' to be Nash. That is, this is the unique case in which T' maximizes its payoff by playing itself. Solving for T' yields the solution of $T' = 0$ (regardless of the values of any of the parameters). Thus, $[0, 0]$, a strategy that never deliberates and always defects, is a best response to itself.

For the strategy $[0, 0]$ to be a Nash, however, it must also be the case that no intuitively cooperative strategy can beat it. This follows straightforwardly. The payoff that strategy $[0, 0]$ gets against itself is 0 (since neither player is paying a cost of cooperation to benefit the other or paying a cost to deliberate). Any intuitively cooperative strategy, on the other hand, is going to incur a cooperation cost c on the fraction of interactions that it cooperates. Moreover, since the $[0, 0]$ agent is always defecting, this intuitively cooperative strategy receives no benefit from the $[0, 0]$ agent. Thus, its payoff is always negative and it cannot invade the $[0, 0]$ strategy under any value of p . As a result, $[0, 0]$ (referred to as the "Intuitive defector (ID)" strategy profile in the main text) is always a Nash equilibrium.

2.3 Purely deliberative equilibrium

Next we investigate the other boundary case of $[0, d]$, a purely deliberative agent that never uses intuition. Note that because this agent never uses intuition, the intuitive response S_i is irrelevant, such that the strategy $[0, d]$ is functionally identical to $[1, d]$. We therefore refer to this strategy as $[-, d]$. To see whether this strategy can be Nash, we start by asking what the best response $[0, T_b]$ is when playing against $[-, d]$. Using the expression above, we find $T_b = p(b - c)$. This makes it seem that $[-, d]$ is not Nash because $T_b \neq d$ (except in the special case where $d = p(b - c)$).

However, because d is the maximum cost of deliberation, T is bounded such that $0 \leq T \leq d$. Therefore, when $d < p(b - c)$, this best response $T_b = p(b - c)$ lies outside the range of possible T values. Recall that because $\frac{\partial^2}{\partial T^2} \pi([0, T], [0, T']) < 0$ is satisfied for all T, T' , the payoff $\pi([0, T], [-, d])$ decreases monotonically as T moves further from the best-response value $p(b - c)$. Thus, when $d < p(b - c)$ (such that the best response is greater than the maximum value of T), the value of T within the allowed interval which best responds to $[-, d]$ is in fact $[-, d]$ itself (i.e. the maximum allowed value of T).

We find a similar result when asking which intuitively cooperative strategy best-responds to $[-, d]$. Solving $\frac{\partial}{\partial T} \pi([1, T], [-, d]) = 0$ gives a best response of $[1, c(1 - p)]$. Thus, by the logic from the preceding paragraph, $[-, d]$ cannot be beaten by any intuitively cooperative strategies if $d < c(1 - p)$. As a result, we see that the purely deliberative strategy $[-, d]$ can be Nash when the maximum cost of deliberation is sufficiently small, such that both $d \leq p(b - c)$ and $d \leq c(1 - p)$ are satisfied.

This result is natural - if deliberating were free, it would obviously be better to deliberate in our model than to use intuition. Thus it is no surprise that there is a minimum d above which it is no longer worth paying to deliberate on all occasions. Given the wide-spread use of intuition by humans, we believe it is a safe assumption that the $d > p(b - c), c(1 - p)$ condition is satisfied.

2.4 Intuitively cooperating equilibria

We next consider whether any intuitively cooperative strategy profile is a Nash. Following the procedure used above, we calculate the expression for the payoff that an intuitively cooperative agent with strategy profile $[1, T]$ gets against an intuitively cooperative agent with strategy profile $[1, T']$:

$$\begin{aligned} \pi([1, T], [1, T']) = & \frac{((1 - p)(-b) - (b - c)p + \frac{T}{2})(\frac{T'}{d} - 1)T}{d} \\ & + (p(b - c) - (1 - p)(c - b))(1 - \frac{T}{d})(1 - \frac{T'}{d}) \\ & - \frac{(\frac{T}{2} - p(b - c))TT'}{d^2} \\ & - \frac{((1 - p)c - p(b - c))(1 - \frac{T}{d})T'}{d}. \end{aligned}$$

We then find the best-response T_b by solving for when the partial derivative of this expression with respect to T is 0, yielding

$$T_b = (1 - p)c.$$

Thus, an intuitively cooperative agent's best response against another intuitively cooperative agent is to deliberate only in cases where the cost of deliberation is not greater than $(1 - p)c$. Note that this is the product of the probability of a 1-shot game occurring $(1 - p)$ and the cost of cooperating c , which is precisely the expected benefit of deliberation for an intuitive cooperator (since what deliberation does here is allow the agent to override her cooperative intuition when she finds herself in a 1-shot game).

In order to test whether the strategy profile $[1, c(1 - p)]$ is Nash, we must also consider whether any intuitively defective strategy profile $[0, T']$ can beat it. To do this, we find the intuitively defective strategy profile that is a best response against the optimal intuitively cooperative strategy profile $[1, (1 - p)c]$ by solving $\frac{\partial}{\partial T} \pi([0, T], [1, (1 - p)c]) = 0$ for T .

This yields the strategy profile $[0, p(b - c)]$ as the intuitively defecting strategy that performs best against the intuitively cooperative strategy $[1, (1 - p)c]$. We then find the conditions under which the optimal intuitively cooperative strategy profile does better against itself than the best response intuitive defecting strategy profile does,

$$\pi([1, (1 - p)c], [1, (1 - p)c]) \geq \pi([0, p(b - c)], [1, (1 - p)c]),$$

in order to find out when $[1, (1 - p)c]$ is Nash. We find that this inequality is satisfied when $p \geq \frac{c}{b}$.

It is also necessary to consider whether $[1, c(1 - p)]$ can be beaten in the boundary case where $d > (1 - p)c$, but $d \leq p(b - c)$, such that the best response against $[1, (1 - p)c]$ is actually $[-, d]$ (as $T = p(b - c)$ is outside the allowed range). Doing so, we find that it is always the case that $\pi([1, (1 - p)c], [1, (1 - p)c]) \geq \pi([-, d], [1, (1 - p)c])$ when $p \geq \frac{c}{b}$. Thus the purely deliberative agent $\pi([-, d])$ cannot invade the intuitively cooperative strategy under these conditions.

We therefore conclude that the intuitively cooperative strategy profile $[1, (1 - p)c]$ is a Nash equilibrium when $p \geq \frac{c}{b}$.

2.5 Summary of Nash results

In sum, we find two main equilibria:

1. The Intuitive Defector (ID) strategy profile that intuitively defects ($S_i = 0$) and never deliberates ($T = 0$) is always Nash (the deliberative strategy variables S_1 and S_r are irrelevant, as this strategy never deliberates).
2. The Dual-process Cooperator (DC) strategy profile that intuitively plays TFT ($S_i = 1$), deliberates when the cost of deliberation is no greater than $T = (1 - p)c$, and deliberatively plays TFT in repeated games ($S_r = 1$) and deliberatively defects in 1-shot games ($S_1 = 0$), is Nash when repeated games are sufficiently common, $p \geq c/b$.

In addition, a purely deliberative strategy that never uses intuition ($T = d$, thus the value of S_i is irrelevant), and deliberately plays TFT in repeated games ($S_r = 1$) and deliberately defects in 1-shot games ($S_1 = 0$), is Nash when the maximum cost of deliberation is sufficiently small, $d \leq c(1 - p)$ and $d \leq p(b - c)$, such that it is always worth paying to deliberate. As this behavior is psychologically unrealistic, we focus our evolutionary analyses on parameter regions where d is large enough to make this strategy not an equilibrium.

2.6 Why is there no equilibrium with $S_i = 0$ and $T > 0$?

A notable feature of our Nash results is the absence of a strategy that intuitively defects but uses deliberation to play TFT when faced with a repeated game. Why can't such a strategy be Nash? The answer is as follows. Unlike in 1-shot games, where it is always beneficial for an agent to defect no matter what the other agent does (because she always avoids paying the cost of cooperation c), the benefit of playing TFT in repeated games depends on coordinating with the other agent. Hence, when two intuitively defecting agents interact and play a repeated game, an agent that pays a cost to deliberate and thereby switch to TFT only benefits from doing so when her partner also deliberates (and thus also plays TFT). As a result, the returns from deliberative cooperation in repeated games for these agents depend not only on the benefit of mutual cooperation $b - c$ and the probability of repeated games p , but also on the probability that the other player deliberates.

Specifically, when two intuitive defectors $[0, T]$ and $[0, T']$ interact, the expected gain from deliberating for the first agent is $\frac{p(b-c)T'}{d}$, the product of the probability of there being a repeated game p , the benefit of mutual cooperation $b - c$, and the probability that the partner also deliberates $\frac{T'}{d}$. As a result, she should be willing to pay a maximum cost of deliberation $T^* = \frac{p(b-c)T'}{d}$ to get this benefit; and indeed, as we saw above, the best response to $[0, T']$ is $[0, \frac{p(b-c)T'}{d}]$. Thus, assuming that one's partner has $T' > 0$, there is always an incentive to deviate by deliberating less ($T^* < T'$). In other words, because of the coordination problem presented by cooperation in repeated games, any nonzero amount of deliberation T' among intuitive defectors is unstable and will be out-performed by intuitive defectors who engage in less deliberation. Therefore, the only equilibrium level of deliberation for a population of intuitive defectors is none at all ($T = 0$). (Or, as discussed above, if the maximum cost of deliberation d is sufficiently low, $d < p(b - c)$, then agents with $T' < d$ will instead be beaten by more deliberative agents with $T^* > T'$, resulting in the equilibrium where agents always deliberate and never use intuition.)

2.7 Nash calculations with assortment

We now consider a version of the game with assortment $a > 0$. In the context of population dynamics, assortment represents non-random mixing, such that with probability $(1 - a)$ a given agent plays with another agent selected at random from the population, whereas with probability a that agents plays with another agent having the same strategy as herself. To incorporate assortment in our Nash calculations, we therefore modify the Nash condition to be

$$\forall y : \pi(x, x) \geq ((1 - a)\pi(y, x) + a\pi(y, y)).$$

We then solve for strategies that are best responses to themselves, in the manner described above. (Note that when $a = 0$, this is exactly equivalent to the above calculations.) Doing so finds that

the ID strategy remains the same when assortment is added ($T = 0$), but that the DC strategy now deliberates with $T = (1 - p)(c - ba)$. Note that a consequence of this is that when $a = c/b$, the DC strategy reaches the boundary case of $[1, 0]$, such that $a \geq c/b$ implies no deliberation by DC, just intuitive cooperation (such that DC stops being an actual dual-process strategy).

3 Risk dominance calculations

3.1 Without assortment

Given that we have identified the game's two Nash equilibria, we are now interested in identifying when one equilibrium or the other will be favored by natural selection. For parameters where ID is the only Nash, it is clearly predicted that evolution will lead to ID. When DC becomes Nash, however, ID also remains Nash. Thus knowing when DC becomes Nash is not enough to know when selection will favor DC.

Risk-dominance, which is a stricter criterion than Nash, has been shown to answer this question: in symmetric 2×2 games such as the one we study, when two symmetric equilibria exist, evolution will favor the risk-dominant equilibrium [1].

One Nash risk-dominates another Nash when the first Nash earns a higher expected payoff than the second Nash when there is a 50% chance of playing against either of the two strategies. Or, in population dynamic terms, the risk-dominant strategy profile is the one that fares better in a population where both are equally common.

We now ask when DC $[1, c(1 - p)]$ risk-dominates ID $[0, 0]$ as a function of p . First, we consider the expected payoffs of these two strategy profiles against themselves and each other:

$$\begin{aligned}
 \pi(ID, ID) &= 0 \\
 \pi(DC, ID) &= -c(1 - p)\left(1 - \frac{(1 - p)c}{d}\right) - \frac{(1 - p)^2 c^2}{2d} \\
 \pi(ID, DC) &= b(1 - p)\left(1 - \frac{(1 - p)c}{d}\right) \\
 \pi(DC, DC) &= ((1 - p)(b - c) + p(b - c))\left(\frac{(1 - p)c}{d} - 1\right)^2 \\
 &\quad - \frac{(1 - p)^2 c^2 (.5(1 - p)c - p(b - c))}{d^2} \\
 &\quad + \frac{(1 - p)c(-b(1 - p) - p(b - c) + .5(1 - p)c)\left(\frac{(1 - p)c}{d} - 1\right)}{d} \\
 &\quad - \frac{(1 - p)c(-c(1 - p) + p(b - c))\left(\frac{(1 - p)c}{d} - 1\right)}{d}.
 \end{aligned}$$

DC risk-dominates ID when

$$\frac{1}{2}\pi(DC, ID) + \frac{1}{2}\pi(DC, DC) > \frac{1}{2}\pi(ID, ID) + \frac{1}{2}\pi(ID, DC).$$

Solving for p in the above equation yields the following condition:

$$p > \frac{\frac{1}{2}(2c^2 - bd - cd + \sqrt{d}\sqrt{-4bc^2 + 4c^3 + b^2d + 2bcd + c^2d})}{c^2}.$$

As we will see below, this value of p successfully captures the transition point we observe in evolutionary dynamics from a population of all ID players to a population of all DC players. For example, for $b = 4$, $c = 1$, and $d = 1$, DC begins to risk dominate ID when $p > .30$ (see Figure 2 of main text). Some other values, which we explore in steady state analyses below, include the following:

b	c	d	p at which DC risk-dominates ID
2	1	1	.62
8	1	1	.14
4	2	1	.50
4	.5	1	.19
4	1	.75	.25
4	1	2	.36

3.2 With assortment

When including assortment $a > 0$, the risk dominance condition for DC becomes

$$a\pi(DC, DC) + (1 - a)(\frac{1}{2}\pi(DC, ID) + \frac{1}{2}\pi(DC, DC)) > \\ a\pi(ID, ID) + (1 - a)(\frac{1}{2}\pi(ID, ID) + \frac{1}{2}\pi(ID, DC))$$

with the DC agent's deliberation threshold now being $T = (1 - p)(c - ba)$ (as shown above in the Nash calculations with assortment).

Thus, the minimum a value at which DC comes to risk dominate ID is given by

$$a > \frac{1}{2(b^2 - 2pb^2 + p^2b^2)} \\ (2bc - 4pbc + 2p^2bc - 2bd + pbd + pcd \\ + \sqrt{(-2bc + 4pbc - 2p^2bc + 2bd + pbd + pcd)^2 - 4(b^2 - 2pb^2 + p^2b^2)(c^2 - 2pc^2 + p^2c^2 + pbd - 2cd + pcd)}).$$

4 Evolutionary dynamics

4.1 Basic setup

We now turn from Nash calculations to evolutionary dynamics. We study the transmission of strategies through an evolutionary process, which can be interpreted either as genetic evolution or as social learning. In both cases, strategies that earn higher payoffs are more likely to spread in the

population, while lower payoff strategies tend to die out. Novel strategies are introduced by mutation in the case of genetic evolution or innovation and experimentation in the case of social learning.

We study a population of N agents evolving via a frequency dependent Moran process with an exponential payoff function [2]. In each generation, one agent is randomly selected to change strategy. With probability u , a mutation occurs and the agent picks a new strategy at random. With probability $(1 - u)$, the agent adopts the strategy of another agent j , who is selected from the population with probability proportional to $e^{w\varphi_j}$, where w is the intensity of selection and φ_j is the expected payoff of agent j when interacting with agents that have the same strategy with probability a , and interacting with agents picked at random from the population with probability $(1 - a)$.

For ease of calculation, our main analyses focus on the limit of low mutation. Later, we also explore higher mutation rates using agent-based simulations, and demonstrate the robustness of our low mutation limit calculations.

4.2 Limit of low mutation calculation method

In the low mutation limit, a mutant either goes to fixation or dies out before another mutant appears. Thus, the population makes transitions between homogeneous states, where all agents use the same strategy. Here the success of a given strategy depends on its ability to invade other strategies, and to resist invasion by other strategies. We use an exact numerical calculation to determine the average frequency of each strategy in the stationary distribution [3, 4, 5].

Let s_i be the frequency of strategy i , with a total of M strategies. We can then assemble a transition matrix between homogeneous states of the system. The transition probability from state i to state j is the product of the probability of a mutant of type j arising ($\frac{1}{M-1}$) and the fixation probability of a single mutant j in a population of i players, $\rho_{i,j}$. The probability of staying in state i is thus $1 - \frac{1}{M-1} \sum_k \rho_{i,k}$, where $\rho_{i,i} = 0$. This transition matrix can then be used to calculate the steady state frequency distribution s^* of strategies:

$$\begin{pmatrix} s_1^* \\ s_2^* \\ \vdots \\ s_M^* \end{pmatrix} = \begin{pmatrix} 1 - \sum_j \frac{\rho_{j,1}}{M-1} & \frac{\rho_{1,2}}{M-1} & \cdots & \frac{\rho_{1,M}}{M-1} \\ \frac{\rho_{2,1}}{M-1} & 1 - \sum_j \frac{\rho_{j,2}}{M-1} & \cdots & \frac{\rho_{2,M}}{M-1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\rho_{M,1}}{M-1} & \frac{\rho_{M,2}}{M-1} & \cdots & 1 - \sum_j \frac{\rho_{j,M}}{M-1} \end{pmatrix} \begin{pmatrix} s_1^* \\ s_2^* \\ \vdots \\ s_M^* \end{pmatrix}$$

The eigenvector corresponding to the largest eigenvalue (1) of this matrix gives the steady state distribution of the stochastic process.

Note that this method requires discretizing the strategy space, such that there is some finite number of strategies M that agents can select. We consider a strategy space in which: (i) agents' cooperation strategies S_i , S_1 , and S_r are limited to being either 0 (never play the cooperative strategy) or 1 (always play the cooperative strategy); and (ii) agents' maximum cost of deliberation T ($0 \leq T \leq d$) that they are willing to pay in order to deliberate is rounded to the nearest $\frac{d}{10}$ (so T is selected from the set $\{0, d/10, 2d/10, \dots, d\}$). Thus, the strategy space consists of a total of $2 * 2 * 2 * 11 = 88$

strategies.

Using the Moran process, the fixation probability $\rho_{B,A}$ (the probability that a single A mutant introduced into a population of B -players will take over) is calculated according to an exponential fitness function. In a population of i A -players and $N - i$ B -players, the fitness of an A -player f_i and B -player g_i are defined as

$$\begin{aligned} f_i &= e^{w(a\pi(A,A)+(1-a)(\frac{i-1}{N-1}\pi(A,A)+\frac{N-i}{N-1}\pi(A,B)))} \\ g_i &= e^{w(a\pi(B,B)+(1-a)(\frac{i}{N-1}\pi(B,A)+\frac{N-i-1}{N-1}\pi(B,B)))} \end{aligned}$$

where $\pi(A, A)$ is the expected payoff of an A -player against an A -player, $\pi(A, B)$ is the expected payoff of an A -player against a B -player, etc.

The fixation probability of a single A -player in a population of B -players can then be calculated as follows:

$$\rho_{B,A} = \frac{1}{1 + \sum_{k=1}^{N-1} \prod_{i=1}^k \frac{g_i}{f_i}}$$

The calculations presented in the main text numerically evaluate this expression for each strategy pair and then solve for the steady state distribution according to the procedure described above. As shown in Figure S1, these evolutionary calculations are in quantitative agreement with the risk-dominance calculations across p and a values shown in the main text Figure 3.

5 Robustness of evolutionary results

5.1 Robustness to parameter variation

Figure S2 shows results of evolutionary steady state calculations for various parameter sets. In each case, we see a qualitatively equivalent pattern to what is observed in the main text Fig 2a: the steady state transitions from intuitive defection $S_i = 0$ with little deliberation $T \approx 0$ when p is small, to intuitive cooperation $S_i = 1$ with substantial deliberation $T \gg 0$ implementing cooperation in repeated games $S_r = 1$ and defection in 1-shot games $S_1 = 0$ when p is sufficiently large. Then, as p increases further, the steady state value of T decreases. (Note that for some parameter values (e.g. Figure S2 panels b, c, and e), when DC becomes risk-dominant, the equilibrium level of T is close to 1 such that agents almost never use intuition, and therefore initially there is little selection pressure on S_i , leading to $S_i \approx 0.5$.)

Quantitatively, the transition from intuitive defection and non-deliberation to intuitive cooperation and deliberation occurs at precisely the value of p where DC begins to risk-dominate ID; and after this point the average value of T matches that of DC, with $T = (1 - p)c$. Thus, these evolutionary calculations show the power of the Nash calculations for characterizing the behavior of our system.

5.2 Robustness to higher mutation rates and probabilistic strategies

We now compare the results of the steady state calculations presented in the main text with agent-based simulations. These simulations use exactly the same evolutionary process as the calculations described above, but relax two simplifying assumptions made in the calculations: the simulations (i) allow agents' probabilities of playing cooperative strategies S_i, S_1 and S_r to take on any value on the interval $[0, 1]$, instead of only allowing 0 or 1 as in the calculations; allow agents' deliberation threshold T to take on any value on the interval $[0, d]$, instead of only allowing discrete values in steps of $d/10$, and (iii) relax the calculation's assumption of vanishingly small mutation and instead use a relatively high mutation rate of $u = 0.05$. For each set of parameters, we conduct 10 simulation runs, each of which lasts 10^7 generations. We then show the value of each of the 4 strategy variables S_i, S_1, S_r , and T , averaged over all generations of all 10 simulation runs (Figure S3 symbols). For comparison, we also show the low mutation limit calculation results (Figure S3 lines). Critically, Figure S3 shows that these agent-based simulations produce very similar results to the calculations. This demonstrates the validity of the calculation, despite its simplifying assumptions.

6 Generalized coordination game analysis

6.1 Setup

The key idea underlying our model is that cooperation sometimes involves a social dilemma (e.g. the 1-shot PD), but other times involves coordination. In our main model, we focus on the infinitely repeated PD as our example of coordination. Doing so, we find that there are two main strategies that can be Nash in this setup: (i) a strategy that intuitively cooperates and sometimes deliberates when the cost of deliberation is less than $T = c(1 - p)$, and (ii) a strategy that always intuitively defects and never deliberates. (We also find that when the maximum possible cost of deliberation d is especially low ($d \leq p(b - c), c(1 - p)$), agents who always deliberate and never use intuition ($T = d$) can also be Nash.)

Here, we demonstrate that these basic results extend to cooperative interactions that involve coordination more generally, rather than being specific to infinitely repeated PDs. To do so, we consider a game where with probability $1 - p$ agents play the 1-shot PD defined above, and with probability p they play a coordination game with the following payoff structure:

	Cooperate	Defect
Cooperate	$A + B$	$A - C$
Defect	$A + B - D$	A

where $A, C \geq 0$ and $B, D > 0$.

This payoff structure has the following features. First, it captures the essence of coordination problems, which is that you cannot improve your payoff by playing something different from the other person (the penalty of not coordinating when the partner defects is captured by $C \geq 0$, and when the other person cooperates by $D \geq 0$). As we are interested in *cooperative* coordination problems, we introduce two additional features: that the cooperative equilibrium is more efficient (higher payoff) than the non-cooperative equilibrium, captured by $B > 0$; and that it requires coordination to achieve the full benefits of this cooperation, such that defecting when the partner cooperates leads to a strictly lower payoff than cooperating when the partner cooperates, $D > 0$ (rather the more

general coordination requirement of just $D \geq 0$). Note that this payoff structure reduces to the infinitely repeated PD using $A = 0$, $B = b - c$, $C = 0$, $D = b - c$.

Using this much more general specification of cooperative coordination problems, we perform a Nash analysis and ask whether (i) we continue to observe the dual process, intuitively cooperative strategy profile that we found using the repeated PD, and whether (ii) an intuitively defecting Nash that sometimes deliberates, which was not observed using the repeated PD, can occur here. We use the same approach described above, in which we focus our Nash analysis on strategies with $S_r = 1$, $S_1 = 0$, and S_i either 0 or 1.

6.2 Intuitively cooperating equilibria

As we did for the repeated PD model, we calculate the best response deliberation threshold with T_b for an intuitively cooperating agent playing against an intuitively cooperating agent with deliberation threshold T' . We find, as before, that the best response is $T_b = c(1 - p)$, regardless of the value of T' (or any of the coordination game parameters). To determine when this strategy $[1, c(1 - p)]$ is Nash, we next consider under what conditions an intuitively defecting agent could beat it. To do so, we find the best response intuitively defecting strategy against $[1, c(1 - p)]$, which we find to be $[0, pD]$ (note that this matches the result from repeated PD model, where the best response was $[0, p(b - c)]$). We find that $\pi([1, c(1 - p)], [1, c(1 - p)]) \geq \pi([p, pD], [1, c(1 - p)])$, such that $[1, c(1 - p)]$ is Nash, when $p \geq \frac{c}{c+D}$ and $d > c(1 - p)$. (Note, again, that this matches the results from the repeated PD model in which the Dual-process Cooperator was an equilibrium when $p \geq c/b$ and $d > c(1 - p)$.)

6.3 Purely deliberative equilibrium

Next, we consider the boundary case that always deliberates, $[-, d]$. Analogous to the results for the repeated PD version, we find that $[-, d]$ is Nash when $d \leq pD$ and $d \leq c(1 - p)$ (as the best response strategy when these conditions are met has $T > d$).

6.4 Intuitively defecting equilibria

Finally, we consider the intuitively defecting case. Unlike in the repeated PD model, we now find that there are two possible intuitively defecting equilibria.

We begin by considering the boundary case $[0, 0]$. We find that the best responding intuitive defector against $[0, 0]$ is $[0, -Cp]$. Because T cannot be negative, this means that among the allowed values of T , $[0, 0]$ is the best response to itself (following the logic explained above in the repeated PD Nash calculations for the purely deliberative equilibrium). Moreover, we find that no intuitively cooperative strategy can ever do better against $[0, 0]$ than $[0, 0]$ does against itself. Thus, as in the repeated PD model, $[0, 0]$, a strategy that never deliberates and always intuitively defects, is always Nash.

Unlike in the repeated PD model, however, our general calculation of the best response deliberation threshold for an intuitive defector playing against $[0, T']$ gives

$$T_b = \frac{p(C+D)}{d}T' - Cp,$$

such that the intuitive defector strategy that best responds to itself is given by $[0, \frac{cDp}{(C+D)p-d}]$. We find that if the maximum cost of deliberation is sufficiently small, such that $d < pD$ and either $d \leq c(1-p)$ or $0 < C \leq \frac{c(1-p)(Dp-d)}{p(d-c(1-p))}$ are satisfied, then this strategy (which has $S_i = 0$ and $T > 0$) is Nash.

Thus, unlike in the repeated PD model, it *is* therefore possible to have a Nash equilibrium that intuitively defects but sometimes uses deliberation to cooperate in the coordination game. Critically, however, this strategy can never be risk-dominant, and therefore is never favored by selection! Whenever $[0, \frac{cDp}{(C+D)p-d}]$ is Nash, there are always two other strategies which are Nash, and both of these other strategies always risk-dominate $[0, \frac{cDp}{(C+D)p-d}]$: $[0, 0]$ is always Nash; when $d < pD$ and $d \leq c(1-p)$, the purely deliberative strategy $[-, d]$ is also Nash; when $d < pD$ and $d > c(1-p)$ but $0 < C \leq \frac{c(1-p)(Dp-d)}{p(d-c(1-p))}$, then $[1, c(1-p)]$ is also Nash. Therefore, as in the repeated PD model, the more general coordination model finds that an intuitively defecting strategy that uses deliberation to cooperate when it is beneficial to do so can never be favored by selection.

6.5 Summary

In sum, the more general social dilemma versus coordination model we have analyzed provides two main conclusions.

1. If the maximum cost of deliberation d is sufficiently large, we observe precisely the same two equilibria observed in the simpler model: (i) an equilibrium that intuitively cooperates and sometimes deliberates $[1, c(1-p)]$, and (ii) an equilibrium that always intuitively defects and never deliberates $[0, 0]$.
2. If the maximum cost of deliberation d is smaller, more complicated equilibria can emerge, such as an equilibrium that intuitively defects and does sometimes deliberate. Crucially, however, this equilibrium is always risk-dominated by another equilibrium, and therefore will never be selected for.

Thus, the conclusions from the repeated PD model hold across all models where agents sometimes play 1-shot PD social dilemmas and other times play cooperative coordination games: selection can favor dual process cooperators, but not dual process defectors.

6.6 Application to repeated PD with finite continuation probability

Our main analyses used the average payoff per round from an infinitely repeated PD between TFT and ALLD for the game with reciprocity. Here, we use the generalized coordination game calculations above to show that these results extend to the more realistic case of total payoff in a repeated PD between TFT and ALLD where after every round, another round occurs with probability δ (such that on average there are $1/(1-\delta)$ rounds per game), yielding the payoff matrix

PD with continuation probability δ

	TFT	ALLD
TFT	$\frac{b-c}{1-\delta}$	$-c$
ALLD	b	0

where $b, c > 0$, $0 < \delta < 1$.

Thus, in terms of the generalized coordination game, this gives $A = 0$, $B = \frac{b-c}{1-\delta}$, $C = c$, $D = \frac{b-c}{1-\delta} - b$. Plugging in these values, we find that the DC strategy continues to be specified by $[1, c(1-p)]$, just as it was for the infinitely repeated PD, and that the condition for DC to be an equilibrium becomes $p \geq \frac{c}{\frac{b-c}{1-\delta} - b + c}$ and $d > c(1-p)$.

6.7 Application to general PD with reciprocal consequences

Finally, we use our generalized results to show that the conclusions of the main model, which used an infinitely repeated PD, extend to the general PD with reciprocity framework outlined in the main text. Here, with probability p , the PD payoff structure is modified such that when one player defects and the other cooperates, the defector's payoff is reduced by α and the cooperators payoff is increased by β , yielding the payoff matrix

PD with Reciprocal Consequences

	C	D
C	$b - c$	$-c + \beta$
D	$b - \gamma$	0

where $b, c, \gamma, \beta > 0$.

In our main model, we focused on the case where $\gamma = b$ and $\beta = c$, yielding a payoff structure that is equivalent to average payoff per round of TFT and ALLD playing an infinitely repeated PD. Plugging this more general form into our results for the cooperative coordination game (using $A = 0, B = b - c, C = c - \beta, D = \gamma - c$), we find that the DC strategy continues to be specified by $[1, c(1-p)]$, just as it was for the infinitely repeated PD, and that the condition for DC to be an equilibrium becomes $p \geq \frac{c}{\gamma}$ and $d > c(1-p)$.

References

- [1] Michihiro Kandori, George J Mailath, and Rafael Rob. Learning, mutation, and long run equilibria in games. *Econometrica: Journal of the Econometric Society*, pages 29–56, 1993.
- [2] Tibor Antal, Martin A Nowak, and Arne Traulsen. Strategy abundance in 2×2 games for arbitrary mutation rates. *Journal of Theoretical Biology*, 257(2):340–344, 2009.
- [3] Drew Fudenberg and Lorens A Imhof. Imitation processes with small mutations. *Journal of Economic Theory*, 131:251–262, 2006.
- [4] Lorens A Imhof, Drew Fudenberg, and Martin A Nowak. Evolutionary cycles of cooperation and defection. *Proceedings of the National Academy of Sciences of the United States of America*, 102(31):10797–10800, 2005.

- [5] David G Rand and Martin A Nowak. The evolution of antisocial punishment in optional public goods games. *Nature Communications*, 2:434, 2011.

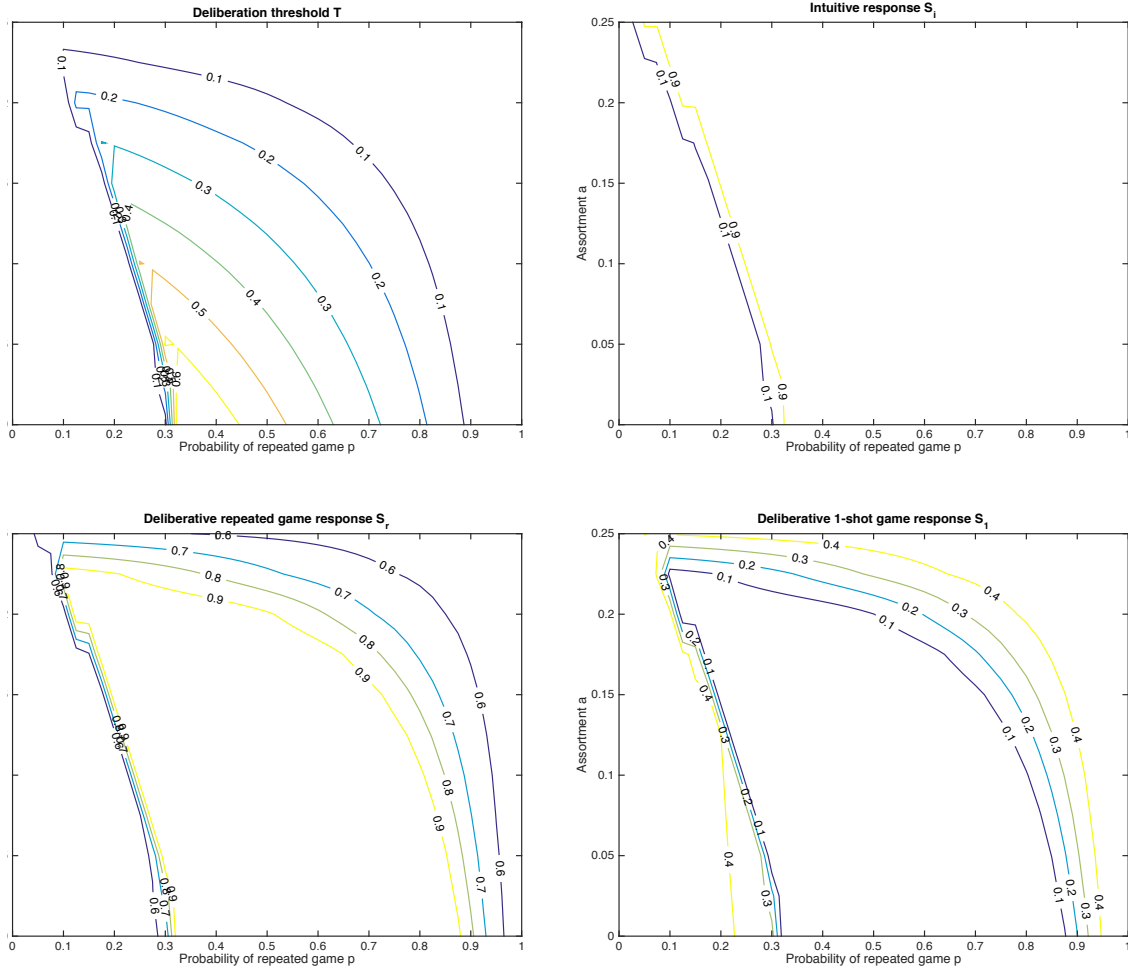


Figure S1: Evolutionary calculations of the steady state distribution using $N = 50$, $b = 4$, $c = 1$, $d = 1$, $w = 6$, for various values of p and a . Shown are the average values of T (a), S_i (b), S_r (c), and S_1 (d). We see quantitative agreement with the risk-dominance calculations shown in the main text Figure 3: S_i is near 0 when ID is risk-dominant and near 1 when DC is risk-dominant; T is near 0 when ID is risk-dominant and equal to $(c - ba)(1 - p)$ when DC is risk-dominant; and S_r is near 1 while S_1 is near 0, except when T is close to zero such that there is little selection pressure on deliberative responses, leading neutral drift to pull S_r and S_1 toward 0.5.

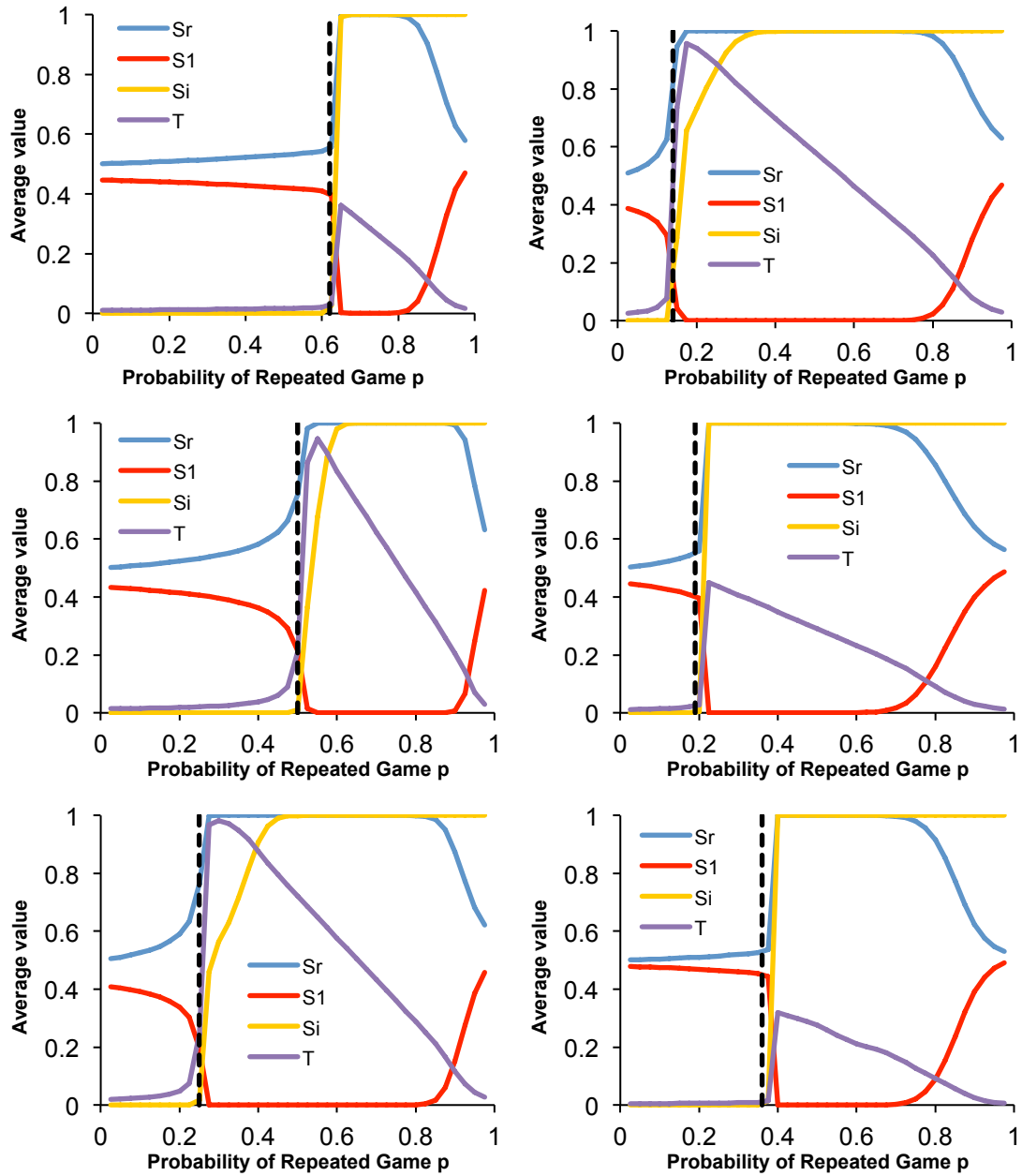


Figure S2: *Evolutionary calculations of the steady state distribution using $N = 50$, $a = 0$ and (a) $b = 2$, $c = 1$, $d = 1$, $w = 6$; (b) $b = 8$, $c = 1$, $d = 1$, $w = 3$; (c) $b = 4$, $c = 2$, $d = 1$, $w = 5$; (d) $b = 4$, $c = .5$, $d = 1$, $w = 6$; (e) $b = 4$, $c = 1$, $d = .75$, $w = 5$; (f) $b = 4$, $c = 1$, $d = 2$, $w = 5$. The point at which DC transitions to risk-dominating ID is presented as a dotted black line for comparison. (Note that because of our use of exponential fitness, for certain parameter sets a smaller selection strength w was needed to prevent the post-exponentiation fitnesses from exceeding MATLAB's computational limits.)*

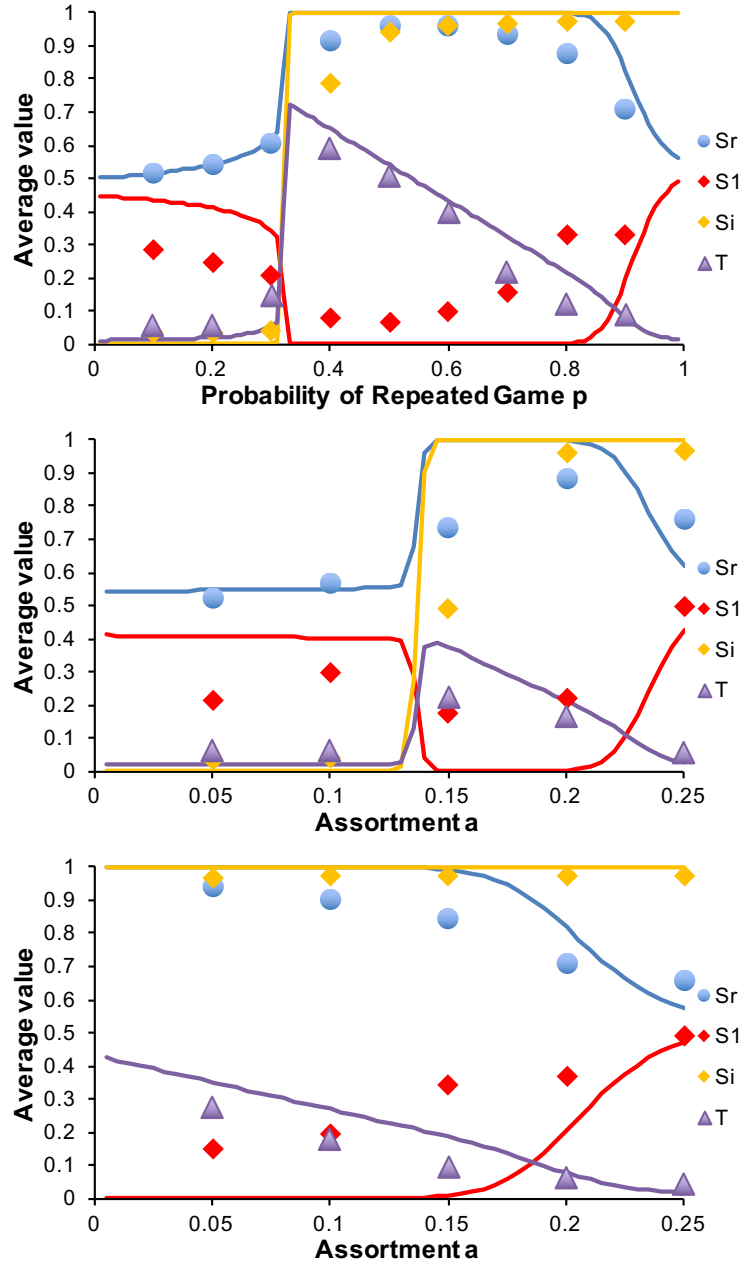


Figure S3: Results of agent-based simulations (symbols) and steady state calculations (lines) showing the average value of each strategy variable, using $N = 50$, $b = 4$, $c = 1$, $d = 1$, $w = 6$. (a) Fixing $a = 0$; (b) fixing $p = 0.2$; (c) fixing $p = 0.6$.