nature human behaviour

LETTERS

PUBLISHED: 7 AUGUST 2017 | VOLUME: 1 | ARTICLE NUMBER: 0152

A general evolutionary framework for the role of intuition and deliberation in cooperation

Stephan Jagau[1,2] and Matthijs van Veelen[1,3]*

In the experimental and theoretical literature on social heuristics, the case has been made for dual-process cooperation. Empirical evidence is thought to be consistent with the idea that people tend to be nice before thinking twice. A recent theoretical paper moreover suggests that this is also the type of dual process one would expect from evolution. In 'Intuition, deliberation, and the evolution of cooperation' by Bear and Rand[1], natural selection never favours agents who use deliberation to override the impulse to defect, while deliberation can be favoured if it serves to undermine cooperation in interactions without future repercussions. Here we show that this conclusion depends on a seemingly innocuous assumption about the distribution of the costs of deliberation, and that with different distributions, dual-process defectors can also evolve. Dual-process defectors intuitively defect, but use deliberation to switch to cooperation when it is in their self-interest to do so (that is, when future repercussions exist). The more general model also shows that there is a variety of strategies that combine intuition and deliberation with Bayesian learning and strategic ignorance. Our results thereby unify and generalize findings from different, seemingly unrelated parts of the literature.

Why humans cooperate is a fascinating question. One could imagine that natural selection would tend to favour selfish individuals over those that confer a benefit to others at a cost to themselves. Yet there is cooperation in a large variety of species, including humans. A number of explanations for the evolution of cooperation have been suggested, and one possible way to classify the different mechanisms is that they fall into three broad categories: population structure, repetition and partner choice. Population structure covers all deviations from random matching, and includes kin selection[2,3], (cultural) group selection[4] and interactions on networks[5,6]. Repetition makes all kinds of reciprocity feasible, which can also stabilize cooperation[7–10]. Finally, picky partners can make each other want to credibly commit to not behaving selfishly, and just not being selfish might be a good way to do that[11–16].

In recent years, the proximate cognitive mechanisms underpinning human cooperation have also received widespread attention. These papers challenge the classical view that cooperative or unselfish behaviour is the result of deliberation and only arises if humans manage to resist their selfish instincts. Instead, experimental psychologists and economists have made the case for dual-process cooperation, where deliberation reins in an intuitive desire to help or cooperate.

A meta-analysis of experimental results[17] finds that in settings where defection dominates cooperation, the average subject cooperates more if it lacks the time or cognitive resources to carefully consider the strategic situation, while there was no effect in settings in which cooperating can be payoff-maximizing, depending on the other player's choice. This supports the intuitive-cooperation pattern[18–21], although others have suggested other possible interpretations of some of the data [22,23] or have proposed alternative explanations based either on a drift-diffusion model of decision making [24] or on the idea that time pressure induces people to make more mistakes, rather than making them behave more cooperatively[25].

A theoretical model[1] moreover suggests that evolution could not have made us any other way—if we are to deliberate at all. The authors find that, depending on parameter values, selection can favour all-out defectors, who never deliberate, or dual-process cooperators, who intuitively cooperate and who sometimes use deliberation to override their cooperative intuition, but it never favours dual-process defectors, who sometimes override their impulse to defect as they deliberate.

In this theoretical model, agents are unsure what game they are in. They might be playing a proper one-shot prisoner's dilemma, without the possibility of future repercussions, or they might be in a situation where the future can in fact cast a shadow over today's actions. The way to think of this second situation is that individuals are really playing a repeated prisoner's dilemma, with a fixed, positive repeat probability. The presence of these possible repetitions leads to a multitude of possible equilibria, both cooperative and uncooperative. In this model[1], these equilibria are collapsed into two possibilities—either all-out defection (AllD) or tit for tat (TFT)—which turns the game effectively from a prisoner's dilemma into a coordination game, where coordination can be on an uncooperative equilibrium or on a cooperative one. The innovative part of their model is that agents have the possibility to resolve the uncertainty about the nature of the interaction and that they can do so at a cost. These costs differ from decision to decision—whereby the cost of the two players are uncorrelated—and in the original model the costs of finding out are uniformly distributed. (There is an earlier model that also formalizes the idea that humans may cooperate in one-shot prisoner's dilemmas, because they are unsure about the nature of the interaction. That is not a model of dual-process psychology, however, because there agents cannot choose to deliberate to learn more about the game.[26])

Here, we construct a general framework, in which individuals can choose to be more or less informed about the details of a game. This information may concern the payoffs of the game or the likelihood of repetitions. Thereby we generalize the previous model[1] in three ways. Our first generalization concerns the distribution of the costs of finding out. If one allows for other possible distributions—such as normal (Gaussian) distributions—we find that dual-process defectors can also evolve, and not just dual-process

[1]CREED, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands. [2]EpiCenter, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands. [3]Program for Evolutionary Dynamics, Harvard University, Cambridge, Massachusetts 02138, USA. *e-mail: C.M.vanVeelen@uva.nl

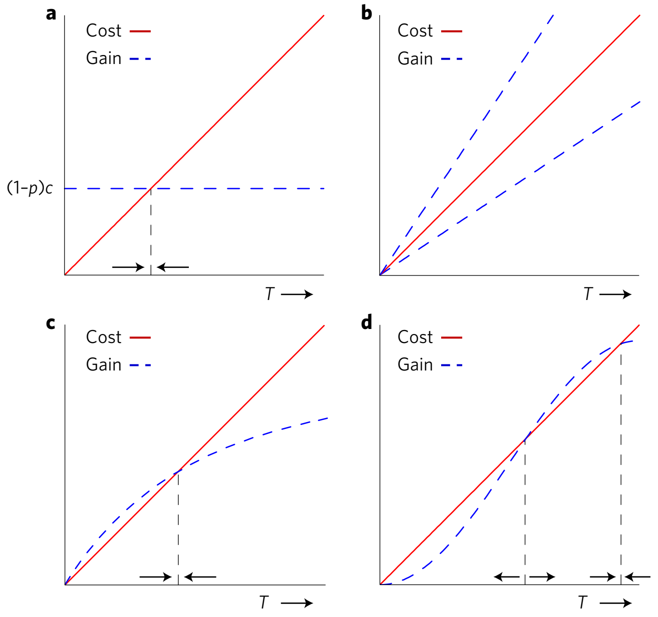

Figure 1 | Dual-process cooperators and dual-process defectors. **a**, Dual-process cooperators. The gain from an additional instance in which the player finds out which game it is does not depend on the threshold. It always equals 1–p, the probability with which the game turns out to be a one-shot prisoner's dilemma, times c, which is the payoff saved by switching from C to D in that game. The costs of getting an additional instance in which the player finds out, however, increases in the current threshold, because the cost is equal to the threshold. **b**–**d**, Dual-process defectors. The benefits of changing from D to C in case the game turns out to be a coordination game, however, depend on what the other player does. What the other player does depends on whether or not he/she also had a draw that made him/her choose to find out. If we assume that the other player uses the same threshold T, then that probability is $F(T)$. The expected benefits of an extra instance of deliberation are therefore $p(b-c)F(T)$. If F is a uniform distribution on $[0,d]$ (**b**), then either $p(b-c) > d$, and this is always worth it (the upper blue dashed line), or $p(b-c) < d$, in which case it is never worth it (the lower blue dashed line). With either full deliberation or no deliberation at all, there is no equilibrium with dual-process defectors. This changes when other distributions for those costs are allowed for. Using $F(z)=1-\frac{1}{(1+z)^4}$ destabilizes the equilibrium at $T=0$, and introduces a stable dual-process defector (**c**). A normal distribution results in two stable equilibria, and an instable one in between (**d**). More detailed computations are in the Supplementary Information.

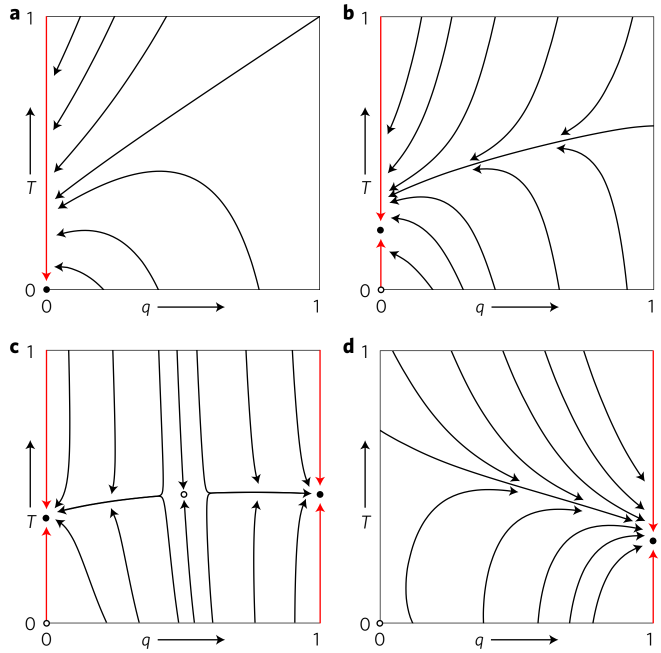

Figure 2 | Adaptive dynamics for different values of *p*. Here we assume that deliberation costs follow $F(z)=1-\frac{1}{(1+z)^4}$. Strategies that evolve have a threshold T that determines when they deliberate, and a probability q with which they cooperate if they choose not to deliberate. **a**, At $p=0.2$ all-out defection is an equilibrium strategy, and it is stable. **b**, At $p=0.4$ there are two equilibria, but all-out defection is not stable, while dual-process defection is. **c**, At $p=0.528$ there are three equilibria within the original strategy set; all-out defection, dual-process defection and dual-process cooperation. Only dual-process defection and dual-process cooperation are stable. With the strategy space extended to allow randomizing between cooperate and defect, there is also an interior equilibrium, which is not stable. **d**, At $p=0.7$ there are two equilibria; all-out defection and dual-process cooperation. Only the latter is stable. Adaptive dynamics with population structure give similar results, where transitions from the situation of one panel to the next happen at lower values of p; see the Supplementary Information.

cooperators or all-out defectors (see Fig. 1). The previous results[1] regarding when equilibria with dual-process cooperators exist turn out to hold in general, as these equilibria do not depend on a particular choice of distribution. Moreover, the dynamics suggest that dual-process defection will typically be stable for low, but not too low, probabilities of the game being repeated, while dual-process cooperation requires higher likelihoods in order to be a stable equilibrium (see Fig. 2). This remains true if we add assortment (see the Supplementary Information).

Our second generalization uncollapses the strategy space, and allows more ways in which the nature of the interaction can vary. For instance, we allow players to be uncertain about whether the probability of a repetition is high or low—making the setup in the previous model[1] a special case, where the low-repeat probability is 0. In the uncollapsed strategy space we find equilibria where players, rather than paying a cost, just wait for time to tell if repeat interactions are sufficiently likely (see Fig. 3). These 'waiting equilibria' exist for a larger set of parameter values than the equilibria with dual-process cooperators from the uncollapsed strategy set.

Our third generalization allows players to condition their future actions on whether or not the other today decides to pay the cost to find out. This introduces the possibility of strategic ignorance, rendering a simple, symmetric version of the cooperate-without-looking equilibria described in Hoffman *et al.*[27]. We show that strategic ignorance can achieve different things. If both players keep themselves in the dark about whether cooperation today is relatively costly or relatively cheap, then that can either allow cooperative equilibria to exist where there were none before, or it can increase efficiency by making players cooperate, not only when cooperation is cheap, but also when it is expensive (see Fig. 4).

Bear, Kagan and Rand[28] have made a different generalization of the previous model[1], and they have also found that there can be equilibria other than all-out defectors and dual-process cooperators. In their generalization, players will get a signal about which of the two games it is, but the signal that the intuitive decision has to work with is less informative than the signal on which the deliberative choice is based. The original model is now a special case, where the intuitive response is based on a totally uninformative signal, and the deliberative response works with a signal that perfectly reveals what the true game is. In this model[28], the authors have shown that this allows for equilibria with dual-process attenders, that sometimes deliberate, and get the better signal, and sometimes do not, in which case they get the worse quality signal, but either way they go with

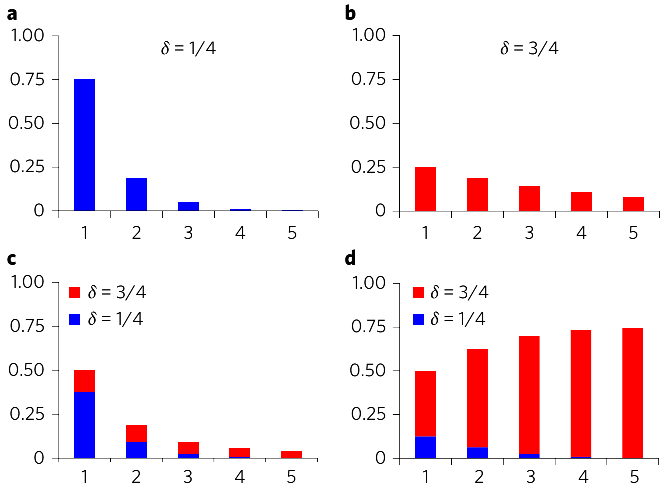

Figure 3 | Repeat probabilities in the repeated prisoner's dilemma with unknown repeat probability δ. a,b, The probability (on the vertical axis) of exactly t repetitions (on the horizontal axis) starts high and decreases quickly for low δ (**a**) and starts low and decreases slowly for high δ (**b**). **c**, In the case that δ could be high or low, both with probability 0.5, the probability of exactly t repetitions starts halfway, goes down relatively quickly at first, but decreases ever more slowly later on. **d**, The probability of reaching round t (on the vertical axis), given that round t–1 is reached (on the horizontal axis), keeps increasing, because if round t–1 is reached, it becomes ever more probable that the true δ is high. Waiting for this probability to be high enough can be part of an equilibrium strategy for the repeated prisoner's dilemma with unknown δ. Note that we start counting at 0, which makes round 1 the 2nd round.

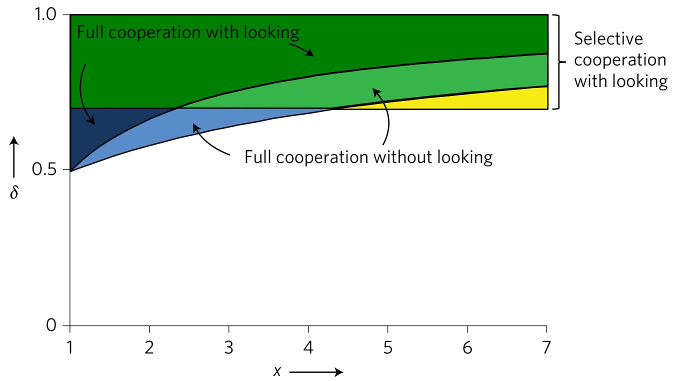

Figure 4 | Strategic ignorance expands the parameter space where full cooperation is possible. The repeat probability δ is on the vertical axis. The temptation to defect in the high temptation game x is on the horizontal axis (it starts at 1, which is the temptation in the low-temptation game). The probability that the high-temptation game is drawn is $p = 0.4$. Full cooperation while looking (without strategic ignorance) is possible above the upper curve, in the dark-blue and dark-green areas. Full cooperation without looking is possible anywhere above the lower curve, in all green and blue areas. Looking and cooperating only in the case that the game with low temptation is drawn is an equilibrium above the straight line, in the yellow and green areas. In the light-blue area, bundling low and high temptation through strategic ignorance is the only way to achieve cooperation at all. In the light-green area, not looking is not needed for cooperation with low temptation, but it does improve the payoffs by bringing cooperation when the temptation is high on board. In the yellow area, the one cooperative equilibrium is to look and cooperate only if the low temptation game is drawn. In the dark-blue area, full cooperation with and without looking are feasible, whereas cooperation only for low temptation is not.

what their signal tells them; they defect if the signal points to the one-shot prisoner's dilemma and cooperate if it points to the coordination game. In this equilibrium, deliberation also increases the chance of cooperation in case the true game is a coordination game, as it does with dual-process defectors. For this equilibrium to exist, the accuracy of the intuitive signal needs to be high enough; with a completely uninformative signal for the intuitive decision, dual-process attenders cannot be selected for.

What was shown in that model[28] therefore is to some extent similar to what we find: in contrast to the original model proposed by Bear and Rand[1], the generalized version allows for equilibria in which deliberation can increase cooperation in case the real game is a coordination game. But our findings also reveal a sensitivity to modelling choices that both previous models[1,28] have in common. Both of these models rely on the distribution of the costs of finding out being uniform, and both use a collapsed strategy space. The fact that the model of ref. 1 is a special case of the model in ref. 28 implies that the sensitivity to these somewhat arbitrary modelling ingredients carries over from the former to the latter. The authors[28] still find that dual-process defectors can never be an equilibrium. By contrast, we show that this is no longer true if we allow for other distributions of the costs. Moreover, for dual-process attenders to reach an equilibrium, the accuracy of the signal on which the intuitive decision is based needs to be sufficiently high, whereas our results imply that even with perfectly inaccurate intuitive signals, one can still get equilibria wih dual-process defectors. Also the uncollapsing of the strategy space in both cases introduces waiting equilibria, with less restrictive equilibrium conditions than the equilibria with dual-process cooperators.

The core assumption of our model, and of the previous models[1,28], is that deliberation makes for individuals that are better (or even fully) informed about the payoffs of the game. That assumption immediately limits the possible predictions of the model concerning the effect deliberation might have. If the actual game is a one-shot prisoner's dilemma, then deliberation can only make individuals cooperate less, or it can have no effect, if their intuition was already to defect; and if the true game is a coordination game, then deliberation can only make individuals cooperate more, or have no effect. Together, the previous models[1,28] and the model described here imply that, other than that, the model itself puts no further restrictions on the possible predictions, and all combinations of the two predictions for each game are possible. All-out defectors would go with deliberation having no effect in either game; dual-process cooperators would have deliberation cause less cooperation in the first game, and have no effect in the second; dual-process defectors would go with no effect in the first game, and more cooperation in the second; and dual-process attenders would go with less cooperation in the first, and more cooperation in the second game. Only additional information about the true value of the different parameters can rule out some possibilities. Moreover, the uncollapsing of the strategy space substantially enlarges the scope for equilibria in which deliberation plays no role.

Methods

Changing the costs of finding out. The model features two one-shot games; Γ_0 and Γ_1.

Γ_0	C	D
C	$b-c$	$-c$
D	b	0

Γ_1	C	D
C	$b-c$	0
D	0	0

With $b > c > 0$, Γ_0 is a prisoner's dilemma; both players playing C yields the socially efficient payoffs, but choosing D maximizes the individual payoff for any choice of the opponent. In Γ_1, on the other hand, choosing action C does maximize individual payoffs, if the other player plays C too. This game can be interpreted as a simplified version of the infinitely repeated prisoner's dilemma, where the only two available strategies are TFT and AllD[1,26]. Two TFT players always cooperate, leading to an average payoff of $b - c$. Any other combination leads to indefinite mutual defection, with average payoffs approaching 0 as the repetitions carry on. The strategies for the two different games go by the same names, because what C and D play in the one-shot prisoners' dilemma coincides with what TFT and AllD play in the first round of the infinitely repeated prisoner's dilemma.

Agents are unsure about which of these games they face. The probability of the game being Γ_1 is p, the probability of the game being Γ_0 is $1 - p$, and the agents know these probabilities. In addition, agents can, at some costs, become completely informed about which game they play. The costs they would have to bear typically vary from decision to decision, are not heritable, and for every decision they are drawn from a distribution F. While in the previous model[1] a uniform distribution of $[0, d]$ is used, we will also allow the use of other distributions.

Strategies for this game need to specify four things: an intuitive decision $x_I \in \{C, D\}$ that determines what the agent does when not informed about the game; two deliberated decisions x_0, $x_1 \in \{C, D\}$ that determine what the agent does when he/she knows he/she is playing Γ_0 and Γ_1, respectively; and the maximum cost $T \in [0, d]$ the agent is willing to pay in order to find out which of the two games it is. A strategy's fitness against another strategy then becomes the expected payoff in the two-player game minus the expected cost of deliberation given the strategy's cost threshold.

Throughout we assume $x_0 = D$ and $x_1 = C$. This assumption does not exclude any strategies that could upset the results. If there is a positive probability that Γ_0 will be played, any strategy with $x_0 = C$ will always do worse than the same strategy with $x_0 = D$. Similarly, a strategy with $x_1 = D$ will never do better than the same strategy with $x_1 = C$. With this assumption, we can denote strategies by the combination of their intuitive response and their deliberation threshold. Equilibria with $x_I = C$ are called intuitively cooperative and those with $x_I = D$ intuitively defecting. Moreover, if the value for the threshold T implies that players sometimes do, and sometimes do not buy information, we refer to these players as dual-process cooperators and dual-process defectors.

Intuitively cooperative equilibria. With costs of finding out that are uniformly distributed, Bear and Rand[1] have shown that the deliberation threshold for dual-process cooperators should be $(1 - p)\ c$. Figure 1a illustrates why this finding is not restricted to costs that are uniformly distributed, but holds for every continuous distribution of costs. If a player wants to find out more often what the true game is, he/she would have to increase his/her threshold, now paying costs also in instances where, before the increase, the costs were above his/her old threshold. Anytime a cost is drawn between the old and the new threshold, he/she pays those costs. Anytime a cost is drawn between the old and the new threshold, he/she also finds out what the game is. If the game turns out to be Γ_1, the intuitive cooperator does not change his/her behaviour. If the game turns out to be Γ_0, the intuitive cooperator changes from C to D, and saves c as a result. Hence the expected marginal benefit from deliberation for an intuitive cooperator is the probability $(1 - p)$ of that happening times c. If those benefits are larger than the costs of the additional case in which he/she found out, the raise in threshold increases his/her payoffs, whether the likelihood of those costs being drawn is small or large. The optimal threshold T^C for dual-process defectors therefore should make those costs and benefits exactly equal, $T^C = (1 - p)\ c$, and this is independent of the distribution of costs.

Besides choosing the optimal threshold, intuitive cooperators must also do better than any possible invader that is an intuitive defector. This requires $p \geq c / b$, which is also independent of the cost distribution (see the Supplementary Information for details). Whether or not there is an equilibrium with dual-process cooperators, therefore, does not depend on the distribution of costs.

Intuitively defecting equilibria. Going from uniform distributions to general ones has no effect on dual-process cooperators. For dual-process defectors, however, there is a marked change. The key to why dual-process defectors are different, is that the benefits of deliberation for intuitive cooperators are independent of what their partner does, whereas the benefits of deliberation for intuitive defectors depend on the choice of their partner. If the game turns out to be Γ_0, nothing changes for the intuitive defector, because both intuition and deliberation call for defection. However, if Γ_1 turns out to be played, which happens with probability p, then deliberation would make the agent play C where the intuitive response was D. As a result, the cooperative benefit $b - c$ can be captured, but only if the opponent cooperates as well. If the other agent is an intuitive defector too, with threshold T', then this means that the probability with which the other one drew a sufficiently low cost to make him/her choose to find out too is $F(T')$, making the expected benefits of deliberation equal to $p\ (b - c)\ F(T')$. Unlike in the case of intuitive cooperators, the distribution of costs does feature in this expression. With T also being the marginal costs of an additional instance of deliberation, T^D is an optimal threshold among intuitive defectors whenever it satisfies $T^D = p\ (b - c)\ F(T^D)$.

With costs drawn from a uniform distribution on $[0, d]$ dual-process defectors that sometimes deliberate and sometimes use their intuition cannot be stable (see Fig. 1b). A uniform distribution makes $F(T) = T / d$ linear, and therefore, in generic cases, it is either best to not deliberate at all, if $p\ (b - c) < d$, or to deliberate all the way, if $p\ (b - c) > d$. This changes if we allow for other cost distributions. While in all cases $T = 0$ remains an equilibrium among intuitive defectors, choosing $F(z) = 1 - \frac{1}{(1+z)^4}$ yields an equilibrium with dual-process defection (Fig. 1c) and choosing a truncated normal distribution creates two new equilibria (Fig. 1d).

Intuitive defectors should not only choose the optimal threshold, but they should also not be outperformed by intuitive cooperators that might try to invade. This implies that the following condition should also be met: $p \leq c / (c + (b - c)\ F(T^D))$ where T^D is the threshold for intuitive defectors. This condition is discussed in more detail in the Supplementary Information.

Dynamics. Amongst equilibria, we would expect that evolution only selects the ones that are stable. The stability of the different equilibria can be examined if we consider a dynamical process. Here we use adaptive dynamics [29–31], which is a standard approach with continuous strategy spaces. As mentioned before, we restrict attention to strategies with sensible deliberate choices; $x_0 = D$, $x_1 = C$. This reduces the strategy space to combinations of an intuitive decision $x_I \in \{C, D\}$ and a threshold $T \in [0, d]$. In order to have well-defined adaptive dynamics, which requires the intuitive decision to be a continuous variable too, we include strategies that randomize their intuitive decision. The intuitive behaviour then becomes a probability q with which cooperation is chosen, which can take values between (and including) 0 and 1.

In order to illustrate how different equilibria can appear and disappear, or become unstable, as the probability that game Γ_1 is drawn changes, we chose $F(z) = 1 - \frac{1}{(1+z)^4}$ as the cost distribution. For this distribution, the gains of deliberation are depicted in Fig. 1c.

For probabilities p at the low end—making it most likely that the game is a proper one-shot prisoner's dilemma—it is a stable equilibrium never to check (set the threshold to $T = 0$) and always to defect (set the probability to cooperate to 0 too; see Fig. 2a). This remains an equilibrium for all p (see Supplementary Information), but it becomes unstable as p increases.

For low intermediate probabilities p, an equilibrium with dual-process defectors appears (Fig. 2b). This equilibrium, with $x_I = 0$ and $0 < T < 1$, is stable. As p increases further, a second stable equilibrium appears, this one with dual-process cooperators ($x_I = 1$ and $0 < T < 1$). Here we also see an unstable equilibrium appear with both $0 < x_I < 1$ and $0 < T < 1$ (Fig. 2c). Finally, for high p we see the dual-process defectors disappear, making dual-process cooperators the only stable equilibrium (Fig. 2d).

This pattern suggests something that also makes intuitive sense. If the probability of future repercussions is low—but not too low to bother—it can be stable to intuitively defect, but reconsider in case you find out that the interaction is repeated. If, on the other hand, there is ample room for the shadow of the future, it can be stable to cooperate by default, and reconsider in case you figure out there is no scope for future repercussions.

Uncollapsing the strategy space. In the baseline model, agents face two types of one-shot interactions: a prisoner's dilemma and a coordination game. The coordination game can be seen as a simplification that captures the essential ingredients of a repeated prisoner's dilemma[1,26]. The two strategies available to the players, called C and D, describe behaviour across both the repeated and the one-shot prisoner's dilemma. Strategy C can be interpreted as TFT, which, in case the game turns out to be one shot, amounts to playing cooperate. Strategy D can be interpreted as AllD, which, in case the game turns out to be one shot, amounts to playing defect. Restricting the set of strategies for the repeated game to TFT and AllD turns it into a coordination game, where players either coordinate on an uncooperative equilibrium of the repeated game, or on a cooperative one. If we uncollapse the strategy space, two things change. Allowing for all ways in which one can play the repeated prisoner's dilemma introduces neutral mutants. These create stepping-stone paths out of equilibrium[9,10], which complicates the dynamics. Compared to the dynamics in Fig. 2, there are many more equilibria, and all of them are left much more easily. And, more importantly, if we uncollapse the strategy space, we also get new kinds of equilibria, where players wait, rather than pay, in order to find out the nature of their interaction. Because such equilibria are interesting in their own right, we will not only allow for more strategies, but also consider richer ways in which one could be uncertain about how likely future repetitions are.

The stage game is the standard prisoner's dilemma:

	C	D
C	$b - c$	$-c$
D	b	0

This game is repeated at an uncertain rate δ. Here we assume that δ can take one of two values, δ_0 or δ_1, and that players know the prior probability p_1 that δ will be δ_1. Also, we assume that $\delta_0 < \frac{c}{b} < \delta_1$. Game Γ_1 as defined in the 'Changing the costs of finding out' section of the Methods now is a special case where $\delta_0 = 0$ and $\delta_1 = 1$. We also go back to the costs of deliberation being uniform, which means that we go back to the setting that, in the uncollapsed strategy set, allows for dual-process cooperators, but not dual-process defectors.

Because the cost–benefit ratio $\frac{c}{b}$ is in between the high and the low value for δ, cooperation could be sustained in case both players knew that the true repeat probability is δ_1, but not if they both knew it is δ_0. Players therefore have an incentive to find out which δ is the true one.

As the game unfolds, players update the probabilities with which they think it is either of the two. The longer the interaction lasts, the more likely it becomes that $\delta = \delta_1$. With Bayesian updating, the posterior probability that $\delta = \delta_1$, given that round $t \geq 0$ has been reached, is given by

$$p_{1|t} = \mathbb{P}(\delta = \delta_1 | t) = \frac{p\delta_1}{(1-p)\delta_0 + p\delta_1}$$

Every extra repetition increases the posterior belief that $\delta = \delta_1$: $p_{1|t+1} > p_{1|t}$ for all $t > 0$, and $\lim_{t\to\infty} p_{1|t} = 1$; see also Fig. 3.

We also uncollapse deliberation by assuming that players can choose themselves in which round they want to buy information. If they make it to that round and buy the information, they pay a cost z, as before, and learn the true value of δ. Other choices for uncollapsing could be used without marked effects on the outcome.

Waiting equilibria. As a benchmark, consider a strategy that we denote by $G(T)$. If this strategy remains uninformed about the true δ, or if it finds out that $\delta = \delta_1$, the strategy $G(T)$ behaves the way 'grim trigger' does in a standard repeated prisoner's dilemma; it starts by cooperating in the first round, and continues to cooperate as long as both players cooperated in all past rounds. After any defection, grim trigger defects forever. However, if the costs are less than T, $G(T)$ buys information before the first round, and always plays D if it turns out that $\delta = \delta_0$.

Combined with a threshold strictly between 0 and 1, this strategy would be a natural equivalent of the dual-process cooperator from before the uncollapsing of the strategy space. As before, with the collapsed strategy space, for $G(T)$ to be an equilibrium, its threshold T should be optimal for a $G(T)$ player against itself. This is the case if it is equal to

$$T = \frac{c - \frac{\delta_0}{1-\delta_0}(b-c)}{\frac{1}{1-p_1} - \frac{\delta_0}{1-\delta_0}(b-c)}$$

Besides getting the threshold right, $G(T)$ should of course also do better than AllD against $G(T)$.

Now consider a strategy $D_kG(T)$, that defects in the first k rounds, cooperates in round k, and then keeps cooperating if the other cooperated too. If the other defects even once after round $k-1$, then this strategy punishes by defecting forever after. Moreover, if round k is reached, this strategy buys information, provided that the costs are below its threshold T, and defects if it turns out that $\delta = \delta_0$.

In a standard repeated game, delaying the onset of cooperation would not matter for whether a strategy is an equilibrium. After k rounds the repeated game looks exactly as it looked at the beginning. In this new setup, where there is uncertainty concerning the repeat probability, this is different. As the rounds pass, it becomes ever more probable that δ is high, which implies that the expected future benefits from mutual cooperation get ever higher. In the Supplementary Information we show that the conditions for $D_kG(T)$ to be an equilibrium are the same as the conditions for $G(T)$ to be an equilibrium, provided that we replace the prior probability p_1 with the posterior probability $p_{1|k}$. Because these probabilities keep increasing, the equilibrium value for T keeps coming down. This implies that the other equilibrium conditions get ever less stringent. If we now go back to the original setting, where $\delta_0 = 0$, we find that in this case updating is extreme and immediate. If the players make it to the second round, then it is certain that the game is not a one-shot game. Uncollapsing the strategy space therefore gives an alternative equilibrium, where players play D in the first round, and grim trigger, or TFT, from the second round onwards. This strategy has less restrictive equilibrium conditions, and has no use for deliberation, as it already knows with certainty what game it plays once it gets to the second round. The collapsing of the strategy space therefore is not innocuous. It ties behaviour in the two games together, while uncollapsing the strategy space also gives players the opportunity to untie them, creating new equilibria that are at least as stable.

Different things strategic ignorance can achieve. Agents in the previous model[1] do not condition their behaviour on whether or not their partner deliberated. Giving players the option to have their actions depend, not only on whether or not their partner played cooperate or defect, but also on whether or not their partner deliberated, can have interesting theoretical implications[27,32,33]. Moreover, it has been demonstrated experimentally that people do actually engage in such conditioning[34]. If we allow strategies to be contingent on deliberation too, then all of the equilibria that we found before, and that now simply do not make use of that option, would remain equilibria. On top of those, we get other equilibria, in which strategic ignorance has a role. In the Supplementary Information we compute such additional equilibria for the setting we have considered so far. If we want to illustrate that strategic ignorance can achieve different things, it is, however, easier if we consider another way to extend the game. This also gives a nice, symmetric version of the cooperate-without-looking equilibria, as first presented by Hoffman *et al.*[27].

Consider a setting where the stage game, and not the repeat probability, is uncertain. Players are matched in pairs to play a repeated game, where each round one of the following two stage games gets played:

Γ_0	C	D
C	1	−1
D	2	0

Γ_1	C	D
C	1	$-x$
D	$1+x$	0

In every round, p is the probability that players play Γ_1 and $1-p$ is the probability that players play Γ_0. We also assume that $x > 1$, so that the temptation to defect is larger in Γ_1 than in Γ_0. Independent of which stage game is played in a given round, the next round is reached with fixed repeat probability δ. In each round, players can deliberate to find out which stage game is played, and the choice to deliberate is observable by the opponent. For ease of exposition, we assume that deliberation costs are zero.

If both players condition their cooperation on the other choosing not to find out which of the two games is being played, they can force each other to remain in the dark. That can have two different effects (see also Fig. 4). They can make the other take the good with the bad, and not cherry-pick so as to cooperate only when temptation is low. If cherry-picking is also an equilibrium, forcing each other not to look increases the level of cooperation, and it has cooperation when temptation is high piggyback on the instances in which temptation is low. The high temptation game still is a prisoner's dilemma, and extending cooperation to those cases therefore still makes both players even better off.

Alternatively, packaging the high and low temptation together can also make players sustain mutual cooperation in cases where even cherry-picking the instances when temptation is low would not constitute an equilibrium. The reason is that if the players only choose those times that the low temptation game is drawn to cooperate, these moments on average get to be fewer and further between than when they cooperate in all instances. Even cooperation in low temptation cases now depends on bundling them together with the instances when temptation is high.

Data availability. Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Received 27 February 2017; accepted 22 June 2017; published 7 August 2017

References

1. Bear, A. & Rand, D. G. Intuition, deliberation and the evolution of cooperation. *Proc. Natl Acad. Sci. USA* **113,** 936–941 (2016).
2. Hamilton, W. D. The genetical evolution of social behaviour. *J. Theor. Biol.* **7,** 1–16 (1964).
3. van Veelen, M., Allen, B., Hoffman, M., Simon, B. & Veller, C. Hamilton's rule. *J. Theor. Biol.* **414,** 176–230 (2017).
4. Richerson, P. J. & Boyd, R. *Not By Genes Alone: How Culture Transformed Human Evolution* (Chicago Univ. Press, 2005).
5. Szabó, G. & Fath, G. Evolutionary games on graphs. *Phys. Rep.* **446,** 97–216 (2007).
6. Allen, B. *et al.* Evolutionary dynamics on any population structure. *Nature* **544,** 227–230 (2017).
7. Axelrod, R. & Hamilton, W. D. The evolution of cooperation. *Science* **211,** 1390–1396 (1981).
8. Bendor, J. & Swistak, P. Types of evolutionary stability and the problem of cooperation. *Proc. Natl Acad. Sci. USA* **92,** 3596–3600 (1995).
9. García, J. & van Veelen, M. In and out of equilibrium I: evolution of strategies in repeated games with discounting. *J. Econ. Theory* **161,** 161–189 (2016).
10. van Veelen, M., García, J., Rand, D. G. & Nowak, M. A. Direct reciprocity in structured populations. *Proc. Natl Acad. Sci. USA* **109,** 9929–9934 (2012).
11. Frank, R. H. *Passions within Reason: The Strategic Role of The Emotions* (W. W. Norton, 1988).
12. McNamara, J. M., Barta, Z., Fromhage, L. & Houston, A. I. The coevolution of choosiness and cooperation. *Nature* **451,** 189–192 (2008).
13. Fujiwara-Greve, T. & Okuno-Fujiwara, M. Voluntarily separable repeated prisoner's dilemma. *Rev. Econ. Stud.* **76,** 993–1021 (2009).
14. Izquierdo, S. S., Izquierdo, L. R. & Vega-Redondo, F. The option to leave: conditional dissociation in the evolution of cooperation. *J. Theor. Biol.* **267,** 76–84 (2010).
15. Baumard, N., André, J.-B. & Sperber, D. A mutualistic approach to morality: the evolution of fairness by partner choice. *Behav. Brain Sci.* **36,** 59–78 (2015).
16. Jordan, J. J., Hoffman, M., Bloom, P. & Rand, D. G. Third-party punishment as a costly signal of trustworthiness. *Nature* **530,** 473–476 (2016).
17. Rand, D. G. Cooperation, fast and slow: meta-analytic evidence for a theory of social heuristics and self-interested deliberation. *Psychol. Sci.* **27,** 1192–1206 (2016).
18. Rubinstein, A. Instinctive and cognitive reasoning: a study of response times. *Econ. J.* **117,** 1243–1259 (2007).
19. Rand, D. G., Greene, J. D. & Nowak, M. A. Spontaneous giving and calculated greed. *Nature* **489,** 427–430 (2012).
20. Zaki, J. & Mitchell, J. P. Intuitive prosociality. *Curr. Dir. Psychol. Sci.* **22,** 466–470 (2013).

21. Rand, D. G. *et al.* Social heuristics shape intuitive cooperation. *Nature Commun.* **5,** 3667 (2014).
22. Myrseth, K. O. R. & Wollbrant, C. E. Cognitive foundations of cooperation revisited: commentary on Rand *et al.* (2012, 2014). *J. Behav. Exp. Econ.* (in the press).
23. Bouwmeester, S. *et al.* Registered replication report: Rand, Greene & Nowak (2012). *Perspect. Psychol. Sci.* **12,** 527–542 (2017).
24. Hutcherson, C. A., Bushong, B. & Rangel, A. A neurocomputational model of altruistic choice and its implications. *Neuron* **87,** 451–462 (2015).
25. Recalde, M. P., Riedl, A. & Vesterlund, L. *Error Prone Inference from Response Time: The Case of Intuitive Generosity in Public-Good Games* (Univ. Pittsburgh, 2015); http://www.econ.pitt.edu/sites/default/files/working_papers/WP%2015-004.pdf
26. Delton, A. W., Krasnow, M. M., Cosmides, L. & Tooby, J. Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proc. Natl Acad. Sci. USA* **108,** 13335–13340 (2011).
27. Hoffman, M., Yoeli, E. & Nowak, M. Cooperate without looking: why we care what people think and not just what they do. *Proc. Natl Acad. Sci. USA* **112,** 1727–1732 (2015).
28. Bear, A., Kagan, A. & Rand, D. G. Co-evolution of cooperation and cognition: the impact of imperfect deliberation and context-sensitive intuition. *Proc. R. Soc. B* **284,** 20162326 (2017).
29. Metz, J., Geritz, S., Meszéna, G., Jacobs, F. & van Heerwaarden, J. *Adaptive Dynamics: A Geometrical Study of the Consequences of Nearly Faithful Reproduction* Vol. 12, 183–231 (Elsevier, 1996).
30. Dieckmann, U. & Law, R. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *J. Math. Biol.* **34,** 579–612 (1996).
31. Dercole, F. & Rinaldi, S. *Analysis of Evolutionary Processes: The Adaptive Dynamics Approach and Its Applications* (Princeton Univ. Press, 2008).
32. Hilbe, C., Hoffman, M. & Nowak, M. Cooperate without looking in a non-repeated game. *Games* **6,** 458–472 (2015).
33. Pérez-Escudero, A., Friedman, J. & Gore, J. Preferential interactions promote blind cooperation and informed defection. *Proc. Natl Acad. Sci. USA* **113,** 13995–14000 (2016).
34. Jordan, J. J., Hoffman, M., Nowak, M. A. & Rand, D. G. Uncalculating cooperation is used to signal trustworthiness. *Proc. Natl Acad. Sci. USA* **113,** 8658–8663 (2016).

Acknowledgements

This paper benefited from the suggestions of audiences at the Amsterdam Brain and Cognition Symposium (Universiteit van Amsterdam), at the Gummersbacher Kolloquium (Heinrich Heine Universität Düsseldorf) and at the UECE Lisbon Meetings in Game Theory and Applications (Instituto Superior de Economia e Gestão, Lisbon). We thank A. Bear and D. Rand for comments. Financial support of the Research Priority Area Behavioral Economics of the University of Amsterdam is gratefully acknowledged. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

S.J. and M.v.V. designed research, performed research and wrote the paper.

Additional information

Supplementary information is available for this paper.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to M.v.V.

How to cite this article: Jagau, S. & van Veelen, M. A general evolutionary framework for the role of intuition and deliberation in cooperation. *Nat. Hum. Behav.* **1**, 0152 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Competing interests

The authors declare no competing interests.