Imperfect Deliberation, Context-sensitive Intuition, and the Evolution of Cooperation

A computational game-theoretic model of the evolution of cooperation in dual-process agents

By Ari Kagan

Yale University

April 22nd, 2016
Thesis Advisor: David Rand
Co-advisor: Adam Bear

Submitted to the faculty of the Program in Cognitive Science in partial fulfillment of the requirements for the degree of Bachelor of Arts

Additional thanks to:

Mark Sheskin
Joshua Knobe
Shelly Kagan
Gina Novick
Elana Kagan
Rebecca Kagan
and my three anonymous peer reviewers

Abstract

If evolution is a selfish process, how can humans have evolved to cooperate? To help answer this question, recent research conducted by Adam Bear and David Rand incorporated a dual-process model of decision making into a computational game-theoretic model of the evolution of cooperation. This model explored the cognition underlying human cooperation: Is deliberative control necessary to restrain cooperative intuitions in an environment where defecting is beneficial? Or does deliberation allow us to cooperate when cooperating may be payoff maximizing, overriding a selfish intuition to defect? The model put forth by Bear & Rand finds two strategies favored by evolution, depending on the prevalence of situations where cooperation can pay off: "Intuitive Defectors" (agents who intuitively defect and never deliberate) and "Dual-process Intuitive Cooperators" (agents with a cooperative intuition who sometimes use deliberation to defect when in a single-shot game). However, this model assumes that deliberation is always perfectly accurate, and that intuition is never context-sensitive. The current research relaxes these assumptions to both test the ecological validity of the original model and to allow for the study of the evolution of cooperation in a wider variety of domains, by introducing the possibility of imperfect deliberation and context-sensitive intuitions. In doing so, I find that the original findings are rather robust, maintaining their validity even when allowing for a moderate amount of inaccurate deliberation or context-sensitive intuition. However, a new strategy favored by evolution is also discovered, "Harmonious Dual-processors," agents who use deliberation and context-sensitive intuitions to work towards the same goal of cooperating or defecting when in an appropriate context, paying to deliberate for greater accuracy rather than to override an opposing intuition.

Introduction

As the phrase "survival of the fittest" might indicate, evolution is typically understood to be a selfish process. The competition to survive is often a zero sum game, with limited resources, mates, or time – when one agent does better another does worse. However, in spite of this fact, evidence of cooperation is clearly present across the natural world, and is a key element of human social behavior. This finding has greatly puzzled evolutionary psychologists. Cooperation – in which an agent undergoes some cost to benefit another agent – seems inherently *unselfish.* How could the *selfish* process of evolution have resulted in the development of a seemingly pro-social behavior like cooperation? It would seem as though selfishness would be the only evolutionarily stable strategy (a strategy that when prevalent in a population cannot be invaded by an initially rare alternative strategy).

The field of evolutionary game theory has given rise to several potential explanations for the evolution of cooperation (Nowak, 2006; Rand & Nowak, 2013). One game frequently used to study this type of situation is the prisoner's dilemma. Because the prisoner's dilemma is widely thought to model a large variety of social interactions, it proves particularly useful in the study of the evolution of cooperation. In a prisoner's dilemma, two agents simultaneously choose to cooperate or defect, without knowing the other player's choice. The payoff matrix of a prisoner's dilemma game has the following structure:

	Agent 2 Cooperate	Agent 2 Defect
Agent 1 Cooperate	3, 3	-1, 4
Agent 1 Defect	4, -1	0, 0

Table 1. Sample Prisoner's Dilemma payoff matrix. The first number in a given cell represents the row player's payoff and the second the column player's payoff. This is just one of infinitely many possible prisoner's dilemma's payoffs, all of which have the payoff structure of A > B > C > D, where in this case A = 4, B = 3, C = 0, and D =-1.

In a prisoner's dilemma, agents choose between defecting (the selfish choice) and cooperating (the collaborative choice). As can be seen in Table 1, the agents collectively do best when both players cooperate; however, each individual agent always does better by defecting, regardless of what the other player does. In game-theoretic terms, the only Nash Equilibrium (a strategy set where neither player can get a higher payoff by changing their own strategy, holding fixed the other player's strategy) is [defect, defect], in spite of the fact that both players get higher payoffs in [cooperate, cooperate]. Consider, for example, two nations at war stockpiling nuclear weapons. Each country individually does best when it keeps its nuclear stockpile, regardless of what the other country does. However, both nations do better when they both eliminate their stockpile (reducing the risk of nuclear war) than when they both keep the nuclear weapons.

Because evolution functions to maximize individual payoffs, in single-shot games (a game in which the players only play each other once), selection will favor defecting. The only way to encourage cooperation in a prisoner's dilemma is to find a mechanism to make cooperation beneficial to the individual player.

Game theorists have determined a number of mechanisms that might serve as ways to incentivize cooperation. For example, in the context of a repeated game, in which agents play multiple prisoners' dilemmas against each other, it can become payoff maximizing for agents to cooperate because current actions may have downstream effects on future payoffs (Trivers, 1971; Axelrod, 1981). Thus, defecting in the current game may give a temporary boost in payoff, but may ultimately have negative repercussions on future payoffs; this mechanism is known as direct reciprocity. Consider bringing food to a friend's party – it is only beneficial to bring fancy food to your friend's party if they will bring fancy food to your parties in the future. But if they will, you both get the benefit of fancy food at both parties by cooperating. If they won't you are better off defecting and bringing cheap food. Reciprocity, as well as other mechanisms, such as reputation (other agents know to defect in future games based on an agent's reputation as a defector) and assortment (non-random mixing of the population, such that cooperative agents are more likely to interact with other cooperators), can make cooperation payoff maximizing in the long-run, and therefore favored by selection (Nowak & Sigmund, 2005; Van Veelen, García, Rand, & Nowak, 2012). Prisoner's dilemmas where the payoff structure has been altered such that cooperation is now payoff maximizing will be referred to as prisoner's dilemmas with reciprocal consequences, and the example of a repeated game with reciprocity will primarily be used in this paper (even though other mechanisms can be just as effective at inducing cooperation).

One example of how direct reciprocity can promote cooperation in a repeated game is the Tit-for-Tat (TFT) strategy profile. When an agent plays TFT in a repeated prisoner's dilemma, the agent begins by cooperating, and then continues by copying what the opposing player did in the preceding round. Through direct reciprocity, this strategy punishes players for defecting. By defecting in future rounds, TFT essentially eliminates the long-run benefits of exploitation (defecting while the other agent cooperates), thus increasing the relative benefit of cooperation. This transforms the prisoner's dilemma's payoff matrix. A new payoff matrix for a prisoner's dilemma game with reciprocal consequences (such as the use of the TFT strategy) has been shown below, displayed as the average payoff per round of a repeated game.

	Agent 2 TFT	**Agent 2 All-D**
Agent 1 TFT	3, 3	0, 0
Agent 1 All-D	0, 0	0, 0

Table 2. Sample prisoner's dilemma with reciprocal consequences. In this case, the matrix has been modeled using the average payoff per round of two agents engaged in an infinitely repeated game, with possible strategy choices of TFT and always defect (All-D). However, a similar payoff matrix could be displayed for any prisoner's dilemma with reciprocal consequences. Once again, the first number displays the payoff for the row player, while the second number displays the payoff for the column player.

When both agents play TFT, they cooperate every round, and thus the average payoff per round is the same as that of a single-shot game in which both agents cooperate. Similarly, if both agents play All-D, the average payoff per round is the same as that of a single-shot game in which both agents defect. However, the interesting shift is the scenario in which one agent plays TFT while the other plays All-D. Here, the average payoff per round is 0 for both agents. In the first round, the TFT player will cooperate and the All-D player will defect, resulting in exploitation. However, in all subsequent rounds both players will defect. When infinitely repeated, the average payoff per round is the same as if both players had defected for every game, eliminating the benefit/cost of exploitation. Contrast this with the payoffs of exploitation in a single shot game – defecting while another agent cooperates results in a much higher payoff in a single shot game. But, because of the reciprocal nature of TFT, doing so in a repeated game provides no payoff benefit. As a result, the incentive to defect is eliminated. In game-theoretic terms, the introduction of the TFT strategy has introduced a second Nash Equilibrium in which both players cooperate. The same thing occurs with other forms of reciprocal consequences.

Computational Modeling

Laboratory experiments, in which subjects play repeated games involving mechanisms to induce cooperation, have aligned with the theoretical analyses, providing preliminary evidence for explanations of how cooperation might have evolved (Rand & Nowak, 2013). But to truly understand the evolution of cooperation, a more formal model is needed.

Computational models of evolutionary dynamics have been used to further explore the evolution of cooperation. The Axelrod tournaments (Axelrod, 1981) is one of the first, most influential instances of computational modeling of prisoner's dilemma scenarios. In the Axelrod tournaments a number of different strategy profiles were pitted against each other in computer simulated tournaments. It was from these tournaments that the TFT strategy was first discovered, emerging victorious against far more complicated strategies.

Since then, in-depth computational models of evolution have been developed in which simulations of agents in an evolutionary environment can be run for millions of generations. In such models, agents have traits representing their strategy profile. As agents face off against each other to play prisoner's dilemma games, different agents employ different strategy profiles. The probability of an agent "reproducing" is determined by their fitness, a calculation of the payoffs they get from these games. This allows the strongest strategies to replicate, mimicking the natural process of evolution. Through such computer simulations, it is possible to test how various strategies would fare in an evolutionary context, given different environmental constraints and different models of behavior. TFT has been demonstrated to be the evolutionarily stable strategy in a wide variety of environments with the use of such computational models.

Typically these computational models have focused on the action chosen by a given agent, namely whether to cooperate or defect. But recent experimental studies have begun to investigate whether examining the cognitive processes behind the decision to cooperate or defect can provide a deeper level of understanding of the evolution of cooperation. A number of questions spring to mind about the relationship between the cognitive process and the final action. How does an agent arrive at the decision to defect or cooperate? Do different agents employ different decision making processes, and do some agents even employ multiple decision making processes at different times? And, perhaps most importantly, does the incorporation of the decision making process into formal models of the evolution of cooperation alter how strategies like TFT and All-D evolve over time?

Dual-process Decision making

Recently, a dual-process model of decision making has been explored in the context of the evolution of cooperation. In dual-process frameworks, two mental processes – intuition and reason – compete to determine the ultimate choices an agent makes (Kahneman, 2003; Evans, 2008). While the concept of a struggle between intuition and reason has been around for millennia, dating at least back to Aristotle, recent psychology research has solidified these concepts into a more formalized model of decision making. In *Thinking Fast and Slow* (2011), for example, Daniel Kahneman explores the relationship between the two different cognitive systems. As he presents it, an automatic, emotional intuition (referred to as System 1) provides an immediate response in most day to day situations. On the other hand, at certain times a slower, rational thought process (System 2) provides a deliberative response. System 1 is both fast and cheap to use, so in many situations the intuitive reaction is *all* that is used. However, System 1 can also be very error prone. As a result, System 2 sometimes overrides the System 1 reaction with a more calculated and accurate response. But using System 2 is also costly, taking time, requiring cognitive resources, and sometimes incurring social costs (Kahneman, 2011; Tomlin et. al, 2015). To get a better understanding of the potential costs of using System 2, consider the comparative difficulty of determining 137 x 212 (a System 2 calculation) and 2 x 2 (a System 1 intuition). Because of such difficulties and costs, System 2 is only used when the benefits of the increased accuracy from correcting the error prone System 1 response outweigh the increased costs of using System 2.

Dual-process models of decision making have been applied to countless fields in recent years, from biology to psychology to economics. A large body of experimental research has recently examined the connection between the evolution of cooperation and dual-process decision making. One common finding in the experimental study of cooperation has been that subjects cooperate far more in single-shot contexts than would be expected by a standard game-theoretic model of decision making. (Delton, Krasnow, Cosmides & Tooby, 2011). Furthermore, using time pressure to force a fast response has been shown to increase the amount of cooperation in these single-shot contexts (Rand, Greene & Nowak, 2012; Rand et al., 2014), suggesting that the choice to cooperate may in fact be an intuitive System 1 response, rather than a deliberated response conditioned on the context (Delton et al., 2011). Additional research has demonstrated that such manipulations of time pressure do *not* affect the level of cooperation in repeated games (Duffy & Smith, 2014), suggesting that unlike in a single-shot game, deliberation may often be aligned with the intuitive response in a repeated game context. Yet, in spite of the many experimental findings suggesting a relationship between dual-process decision making and the evolution of cooperation, the dual-process framework had not been incorporated into formal computational models of the evolution of cooperation until quite recently.

Intuition, Deliberation, and the Evolution of Cooperation

Adam Bear and David Rand (2016) applied a dual-process framework to the evolution of cooperation, creating a computational model of evolution that places agents in a combination of both single-shot and repeated prisoner's dilemma games. Agents select their strategic response either by responding intuitively (being insensitive to game type, single-shot versus repeated, when choosing whether to cooperate or defect), or by deliberating (paying a cost to identify the current game type, and conditioning their choice of whether to cooperate or defect on game type).

Different game types can have different payoff maximizing responses, so agents may benefit from conditioning their response on the current game type. When an agent plays in a repeated game, strategies such as TFT[1] make it so that agents benefit from cooperation (so long as the other person also cooperates), as discussed above (shown in Table 2). But, when an agent plays in a single-shot game reciprocity holds no power to induce cooperation, as injured agents have no future games to retaliate in - recall that the only Nash Equilibrium in a single shot game is [defect, defect], as seen in Table 1. Agents maximize payoffs by defecting: they will not suffer reciprocal consequences for doing so.

If an agent always uses the same intuitive response regardless of game type, they sacrifice potential payoffs they could gain by conditioning their response on the game type. If, instead, agents stop to deliberate and identify the current game type, they can determine the payoff maximizing response (defect in a single-shot game, cooperate in a repeated game – see Table 1 and Table 2). However, while deliberation allows agents to select the payoff maximizing response, deliberation also entails a cost in real life, taking up time, inflicting social costs, or even simply using limited cognitive resources. Thus, the model was designed such that the benefit of deliberation is *sometimes* outweighed by the cost of deliberation (which varies for each interaction),

---

[1] As mentioned before, and as Bear & Rand discuss in detail, while reciprocity is one possible mechanism to induce cooperation, the payoff structure does not necessarily make use of TFT - other mechanisms can be substituted into the description without any substantive modifications to the model, so long as they impose downstream consequences on defecting in a repeated game. However, the current research will refer to the model as though it employs TFT for the sake of conceptual simplicity.

making it more beneficial in some cases for the agent to use the context insensitive intuition than to pay the sometimes high cost of deliberation.

Bear & Rand's study found two strategies favored by evolution, depending on the environment: Intuitive Defectors (ID), who never cooperate or deliberate, and Dual-process Intuitive Cooperators (DP-IC), who intuitively cooperate but deliberate when the cost is low enough and thereby sometimes defect in single-shot games. The central finding of the study was that there are no Dual-process Intuitive Defectors (DP-ID), agents who intuitively defect but then use deliberation to cooperate in repeated games. Instead, deliberation only develops as a tool to allow for defection in single-shot games (in the case of the Dual-process Intuitive Cooperators). [2]

Ecological Validity

While Bear & Rand created a very reasonable initial model of the evolution of cooperation in dual-process agents, their model makes a number of assumptions that may affect their results. As noted in the discussion section of their paper, the model assumes that agents play in only two game types (single-shot and infinitely repeated games), that the cost of deliberation is sampled from a uniform distribution, that agents cannot condition their responses upon the cognitive styles of their partners or on past behaviors, that deliberation is perfectly accurate, and that intuition is completely context insensitive. While any number of these assumptions could be fascinating to explore in greater detail, the current research focuses on these last two assumptions: perfect deliberation and context-insensitive intuition.

Bear & Rand assume, for the sake of their model, that once an agent pays the cost and deliberates, they are able to identify the game type with perfect accuracy and play the

[2] These findings will be discussed in greater detail later in the paper.

appropriate response 100% of the time. This is a reasonable assumption for a first model to make, but one can easily see that this might not be the most accurate representation of how deliberation takes place in the real world. Instead, when an agent attempts to identify whether they are in a repeated game scenario or in a single-shot interaction in the real world, they may be able to correctly read the situation, but they also might misidentify the game type, incorrectly thinking that it is a single-shot game when it is actually a repeated game, or vice versa. While System 2 deliberation is certainly more accurate than System 1 intuition, System 2 deliberation still does not ensure perfect accuracy in the real world. Agents rarely have perfect information, nor do they always know how to interpret the information they do have. For example, when trying to decide whether to help someone, you may predict that you will not see them again (a single-shot game) and choose to defect, but you might predict incorrectly! Indeed, in a stochastic world, it might be most accurate to say that a single-shot game is not even determined as such with certainty until future events preclude all possibility of a repeated interaction. Thus, it may never be possible for agents to perfectly determine the game type. (Delton et al., 2011).

Bear & Rand's model also assumes that when an agent uses intuition, the intuition is perfectly *in*sensitive to game type, such that an agent using the intuitive response will always play the same response whether in a single-shot game or a repeated game. However, in the real world our intuitions are not perfectly insensitive. Instead, we often have different intuitions when in different contexts. Our instinct to help someone or not may partially depend on whether we know them or on a number of other contextual variables. For instance, you likely do not need to deliberate to determine the game type when interacting with a family member or a very close friend – you simply have a gut intuition to cooperate. On the other hand, you may have a gut intuition to defect when interacting with a stranger while on vacation. Thus, it seems agents may sometimes be able to identify the game type without needing to pay the cost and deliberate. In other words, while System 2 deliberation is more context-sensitive than System 1 intuition, System 1 intuition may nevertheless be context-sensitive at times (Gigerenzer, 1999).

Current Research

In this paper, I aim to examine the ecological validity of the model proposed by Bear & Rand by relaxing the two assumptions described above:

1) **Imperfect deliberation**: I introduce the possibility that an agent will pay the cost to deliberate but then misidentify the game type and play the incorrect response.

2) **Context-sensitive intuitions**: I introduce the possibility that an agent will correctly identify the game type using only their intuitive response, thus avoiding the need to pay the cost of deliberation.

In the following studies, I incorporate these two modifications into the original model proposed by Bear & Rand. In Study 1 I vary the accuracy of deliberation and examine its effects on the strategies favored by evolution. In Study 2 I vary the accuracy of intuition and examine its effects on the strategies favored by evolution. Finally, I discuss the implications that imperfect deliberation and context-sensitive intuitions have on Bear & Rand's original findings, and consider new insights determined by the current research.

Technical Background: Bear & Rand's Original Study

The present studies duplicated the original model developed by Bear & Rand with the exception of the two modifications discussed above. The following section will provide an overview of the original model used by Bear & Rand, and later sections will outline the two precise modifications introduced in the current studies.

Methods

The original model was different from most models of the evolution of cooperation in two dimensions: agents engage in both single-shot games and repeated games in any given generation, and agents can use deliberation to identify an appropriate response for the given game type, rather than always using the same, fixed response.

Game type. In the original model, agents engage in interactions in which they play both single-shot prisoner's dilemmas (PD) and repeated game PDs. The game type for any given interaction is randomized, single-shot games occurring with probability 1-p and repeated games occurring with probability p. The single-shot game is a typical PD, in which each agent can choose to cooperate, in which case the agent pays a cost C to provide a benefit B to the other player, or choose to defect, in which case the agent pays no cost and provides no benefit to the other player. In this case, B was set to 4 and C was set to 1. The generic payoff matrix is displayed below.

	Agent 2 Cooperate	**Agent 2 Defect**
Agent 1 Cooperate	B-C, B-C	-C, B
Agent 1 Defect	B, -C	0, 0

Table 3. The payoff matrix of a single-shot prisoner's dilemma. B represents the benefit a player receives when their partner cooperates. C represents the cost a player pays when they cooperate. Note that this is the same payoff structure as shown in Table 1, as long as B > C > 0.

The notion of a repeated game is operationalized in a slightly more complicated fashion. Since the model cannot have agents play infinitely repeated PDs, the payoffs structure of the repeated game represents the average payoff per round of an infinitely repeated PD, as done previously in Table 2. Once again, because TFT reduces the benefit of exploitation, the respective benefit and cost of exploitation and being exploited are reduced to zero. This transformed payoff structure for the repeated game turns the PD into a coordination game, making it beneficial for an agent to cooperate if and only if their partner also cooperates. The payoff matrix is displayed below.

	Agent 2 Cooperate	**Agent 2 Defect**
Agent 1 Cooperate	B-C, B-C	0,0
Agent 1 Defect	0,0	0, 0

Table 4. The payoff matrix of a repeated game prisoner's dilemma. Once again, B represents the benefit a player recieves when their partner cooperates, and C represents the cost a player pays when they cooperate. Note that this is the same payoff structre as shown in Table 2, as long as $B > C > 0$.

Deliberation. The more interesting feature of the original model is the introduction of the possibility of deliberation into the traditional prisoner's dilemma scenario. The agents are given the ability to reach a decision either via intuition or via deliberation. As discussed before (and as is standard in dual-process models), use of the deliberative response is more costly than intuition; however, the intuitive response is insensitive to the details of the context in which the decision is made.

While deliberation allows the agent to more precisely identify the appropriate response, it comes at a cost, as discussed. The cost of deliberation can vary with context; as a result, the cost of deliberation in any given interaction, d*, is sampled over a uniform distribution between 0 and D. Thus D is an environment variable representing the maximum possible cost of deliberation (the baseline model of Bear & Rand sets D = 1, and as a result d* was sampled from 0 to 1; I follow the same convention). The cost d* is subtracted from the final payoff the agent receives from their decision to cooperate or defect.

Whether an agent deliberates or uses the intuitive response depends upon the relation between this cost, d*, and the agent's deliberation cost threshold (T) which evolves (along with the agent's other responses described below). T represents the highest cost that the agent would be willing to pay to deliberate. Thus, if d* is greater than T the agent will not deliberate. If d* is less than T the agent will deliberate.

In addition to T, an agent's strategy profile specifies their cooperative or non-cooperative responses in each possible scenario. All of the responses used by the model are implemented merely as probabilities: the probability that an agent will cooperate when using this given response. Because the only factors influencing a response are whether the agent is deciding intuitively or, if deliberating, whether the agent believes the current game to be repeated or single-shot, each response variable is simply the probability of cooperating in the specified circumstances. As evolution acts on these responses, the probability of cooperation rises or falls to a stable value.

If an agent does not deliberate (because d* is greater than T), they use their intuitive response, SI, regardless of the game type (as intuition here is context insensitive). However, when an agent uses deliberation (because d* is less than T), the agent perfectly identifies the current game type, repeated or single-shot (as deliberation here is perfectly accurate), and selects the appropriate strategic response (SDR or SD1, respectively). Thus, the deliberating agent will cooperate with probability SDR when in a repeated game and will cooperate with probability SD1 when in a single-shot game.

In summary, the strategy profile of each agent is composed of four variables: SI (the probability that an agent will cooperate when if using intuition), SD1 (the probability that an agent will cooperate if deliberating in a single-shot game), SDR (the probability that an agent will cooperate if deliberating in a repeated game), and T (the highest cost the agent would be willing to pay to deliberate). These responses combine to create the strategy profile played by each agent.

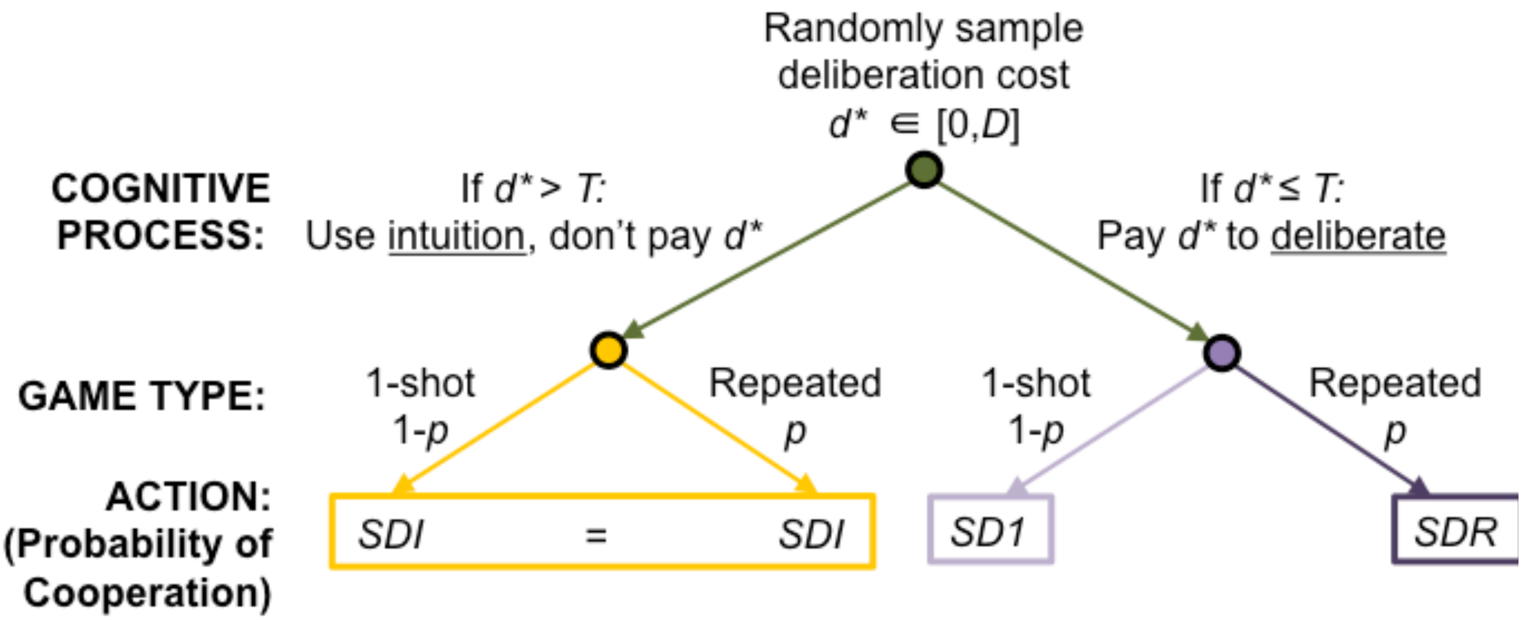

Figure 1. The strategy space of the original model. The four variables, SDI, SD1, SDR, and T are visualized here, along with the sequence of events that takes place in each interaction. First, the agent's cost of deliberation for this interaction d* is sampled uniformly from the interval [0, D]. The 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 the agent bases the cooperation decision the generalized intuitive response SI. Since intuition cannot differentiate between game types in the original model, the agent plays the cooperative response with probability SI, regardless of whether the game is single-shot (probability 1-p) or repeated (probability p). On the other hand, if d* ≤ T, deliberation is not too costly, so the agent pays the cost d* and uses deliberation to condition the response on game type: If the game is single-shot, the agent cooperates with probability SD1, and if the game is repeated, the agent cooperates with probability SDR (Bear & Rand, 2016).

Simulation. The program simulation is designed and run in the MATLAB programming environment. The simulation is carried out over a population size N (50) using the Moran process, one of several popular mechanisms used to represent evolution. In the Moran process, the fitness of each agent determines the probability that it will "reproduce" (Moran, 1958). In each generation one of the N agents is randomly selected to be removed from the population and "die." From the remaining N-1 agents, one agent is selected to be duplicated, with the probability of any agent being selected proportional to the evolutionary fitness (see below) of that agent. With some probability, μ, a mutation occurs, and an agent with a random strategy profile is created. Otherwise, the chosen agent is copied. In this model, the agents then engage in a new round of PDs to determine their updated payoffs. Each agent plays against every other agent once, thus playing N games with each game type[3] determined by p, and each game with a different randomly sampled d*. At the end of the round, an agent's evolutionary fitness is determined by the sum total payoff received from playing every other agent.

In an *agent-based simulation*, this entire process is repeated for 10 million generations, by which time the strategy favored by evolution has typically arisen within the population. On the other hand, a *steady-state calculation* focuses on an environment with a lower mutation rate such that when a mutant arises in the population, the mutation either quickly dies off or takes over the population before another mutation occurs. As a result, the population goes through a series of transitions between homogenous states in which the entire population uses a single strategy profile. This allows us to use a precise numerical calculation, the steady-state calculation, to determine the fraction of time the population would spend in each state if the actual agent-based simulations were run for an infinite period of time. It is this steady-state calculation that is used to generate the figures provided in this paper (Bear & Rand, 2016).

Results

As mentioned previously, Bear & Rand found two strategies favored by evolution. When $p < 0.3$, the dominant strategy (the strategy employed by more than 50% of the population) is the Intuitive Defector, an agent which has an intuitive response of always defecting (SI = 0) and never deliberates (T = 0). On the other hand, when $p > 0.3$, the dominant strategy is the Dual-process Intuitive Cooperator, an agent which has an intuitive response of always cooperating (SI = 1) but sometimes deliberates (T > 0). The DP-IC typically cooperates, but deliberation

[3] Remembering, of course, that both game types are in fact implemented as single-shot games with different payoff structures.

sometimes causes the agent to defect in a single-shot game. DP-IC becomes the dominant strategy at some critical threshold, when DP-IC "risk dominates"[4] ID.

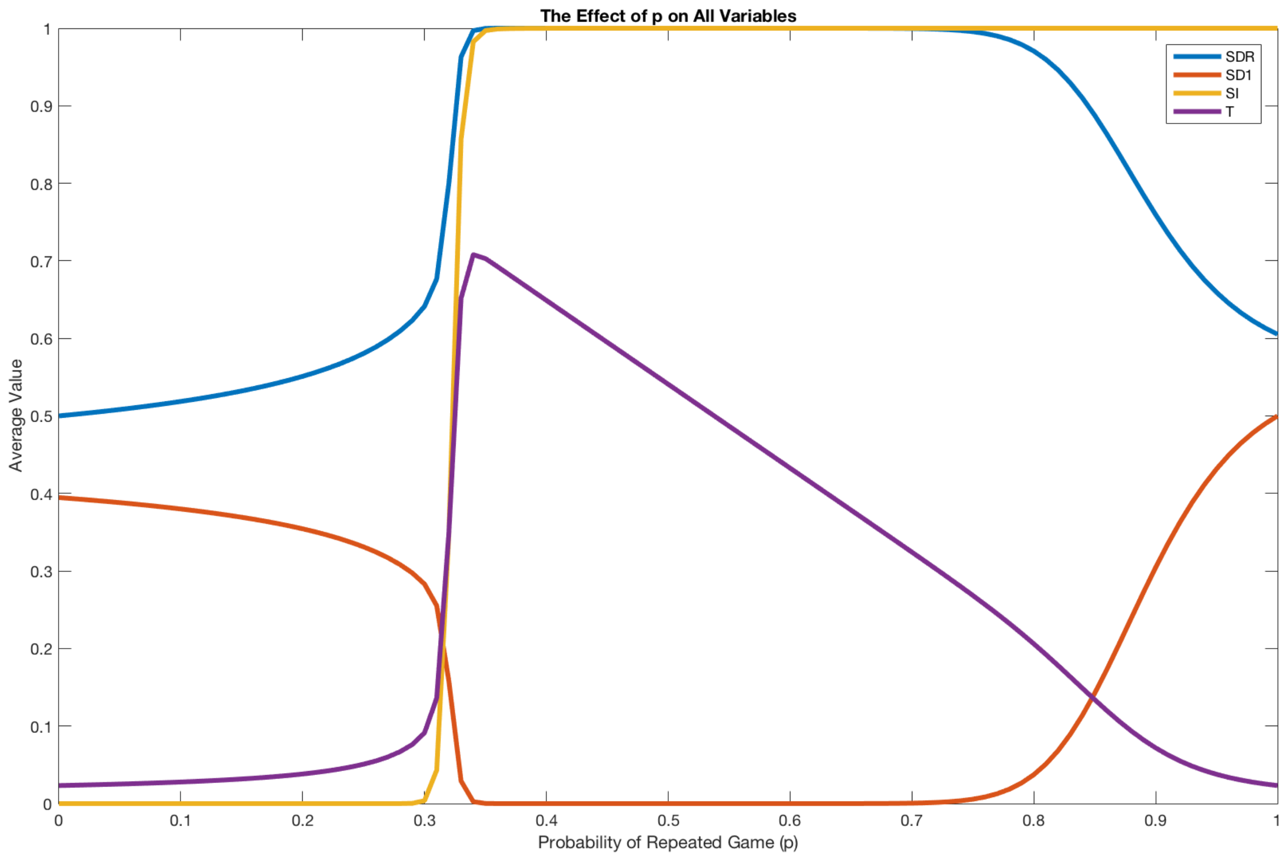

Figure 2. The effect of p on all variables. B = 4, C = 1, D = 1. The X axis is the environment variable p, the probability of being in a repeated game. The Y axis is systematically ambiguous, as follows: shown in purple is T, the maximum deliberation cost the agent is willing to pay (between 0 and 1, as D = 1); in blue is SDR, the agent's probability of cooperating when deliberating and believing they are in a repeated game; in red is SD1, the agent's probability of cooperating when deliberating and believing they are in a single-shot game; in yellow is SI, the agent's probability of cooperating when responding intuitively. The figure was generated using a steady state calculation – the fact that deliberation is not perfectly at floor level when $p < 0.3$ is the result of noise, and would not display in a Nash Equilibrium calculation.

Fig. 2 shows that when $p < 0.3$, T is 0, representing a lack of any deliberation in the population. SI is also at 0, representing an intuitive response to defect. Because deliberation is at floor level, the deliberative responses, SDR and SD1, are never used: their values do not affect

[4] This means that an agent placed in a population with half of the agents playing ID and half of the agents playing DP-IC would get a higher payoff by playing DP-IC than by playing ID.

fitness, so there is little selection pressure on their values. Random variation (referred to as neutral drift) caused by the possibility of random mutation built into the model pulls SDR and SD1 towards 0.5.[5] Finally, because p is small, most interactions are single-shot games, and as a result it is most beneficial for the intuitive response to be defect. With no deliberation, and an intuition to defect, this strategy profile is ID.

Once $p > 0.3$, several changes occur. T quickly climbs much higher, representing the introduction of deliberation into the population. Additionally, SI switches to 1, representing an intuitive response to cooperate. Additionally because deliberation now takes place, there is now selection pressure on the deliberative responses, and they separate into SDR = 1 (cooperation in a repeated game after deliberation) and SD1 = 0 (defection in a single-shot game after deliberation). Because p is now larger, repeated games are more common and it becomes beneficial for cooperation to be the intuitive response. This strategy profile is DP-IC, an agent who will intuitively cooperate, but when deliberates and identifies a single-shot game, will defect.

As the p value continues to approach 1, T declines once more to 0, representing a reduction in the amount of deliberation in the population. As T approaches floor level once more, neutral drift once again pulls SD1 towards 0.5.[6] As the probability of being in a repeated game nears 1, the probability of deliberation identifying a single-shot game and switching the response to defect, thus increasing the agent's payoff, decreases. As a result, deliberation is used less and less.

---

[5] Because deliberation is not perfectly at floor level, there is still very slight selection pressure on SDR and SD1. This is not strong enough to significantly separate the two deliberative responses. However, it does pull the deliberative responses slightly below 0.5 on average. With so little difference between the two responses, they act somewhat like a context-insensitive intuitive response. Thus, when $p < 0.3$, and the intuition is to defect (SI = 0), the deliberative responses get pulled slightly below 0.5 towards 0.

[6] However, since SI now equals 1, the slight levels of deliberation pull the deliberative responses slightly above 0.5 (instead of slightly below as when SI = 0).

Why is there no deliberation when $p < 0.3$ and the intuitive response is to defect? One might think it would be beneficial to deliberate and switch to cooperation when in a repeated game, resulting in a strategy profile we might refer to as a Dual-process Intuitive Defector (DP-ID). The key to understanding why there are no DP-IDs is to first understand why there are DP-ICs.

Why isn't the dominant strategy when $p > 0.3$ the Intuitive Cooperator (IC), an agent which has the intuition to cooperate, but never deliberates? In a world where IC is the dominant strategy, the majority of agents intuitively cooperate. If an IC agent stops to deliberate and identifies the current game as single-shot, the agent could defect and get a higher payoff (provided a low enough d*) than if the agent were to have intuitively cooperated. When the higher payoff from switching to defection in a single-shot game sometimes outweighs the cost of deliberation, deliberating can be payoff maximizing. Thus, DP-IC arises as the dominant strategy, with T set such that the maximum cost DP-IC will pay to deliberate is the expected gain from switching to defection in one-shot games ($T = C * [1 - p]$).

On the other hand, an ID agent exists in a population where the majority of agents intuitively defect. If the agent were to stop to deliberate, identify a repeated game, and switch to cooperation, the agent *only* would receive a higher payoff if the other agent *also* were to switch to cooperation – which requires the other agent to also deliberate and realize it is in a repeated game. As a result, the benefit of switching to cooperate is discounted by the extent to which the other agent fails to deliberate, leading each agent to want to deliberate somewhat less frequently than their partner. As a result, the population doesn't stabilize until deliberation hits floor level at $T = 0$. Therefore, deliberation does not take place when the intuitive response to defect, and DP-ID is not a dominant strategy.

Because the benefits of defecting require only one agent but the benefits of cooperation require both, occasionally defecting in a world of cooperators can be payoff maximizing in ways that occasionally cooperating in a world of defectors can never be. As a result, DP-ID is not a strategy favored by evolution.

Study 1: Imperfect Deliberation

Methods

In this first modification of the model, I introduce the possibility of imperfect deliberation. The possibility of imperfect deliberation was operationalized through the use of a new environment variable, AD, which represents the accuracy of deliberation. As in the original model, the agent deliberates and pays the associated cost d*, so long as d* is below the agent's deliberation threshold, T. However, in this modified model, after paying the deliberation cost the agent misidentifies the game type with some probability (1 - AD) and plays the response for the opposite game type as a result. For instance, if an agent is actually in a single-shot game, deliberates and misidentifies the situation, the agent would then cooperate with probability SDR rather than playing the appropriate response of cooperating with probability SD1. AD was varied as the independent variable in Study 1 along the spectrum from 1 to 0.5. When AD is set to 1, deliberation is successful 100% of the time (which is the perfectly accurate deliberation seen in the original Bear & Rand model). However, as AD decreases to 0.5, the accuracy of deliberation decreases, such that at 0.5 the agent properly identifies the correct game type 50% of the time, or at chance. Note that at AD = 0 the agent always perfectly fails to identify the game type, but deliberation will be perfectly informative, nevertheless. Because the agent identifies the game type 100% inaccurately, agents will eventually evolve to use the deliberative responses in the opposite contexts of what their name would suggest. SDR and SD1 would now be played in single-shot and repeated games, respectively, but evolve accordingly to be defect and cooperate. Because of this I represent AD on a 1 to 0.5 scale, as 0 to 0.5 would be duplicative.

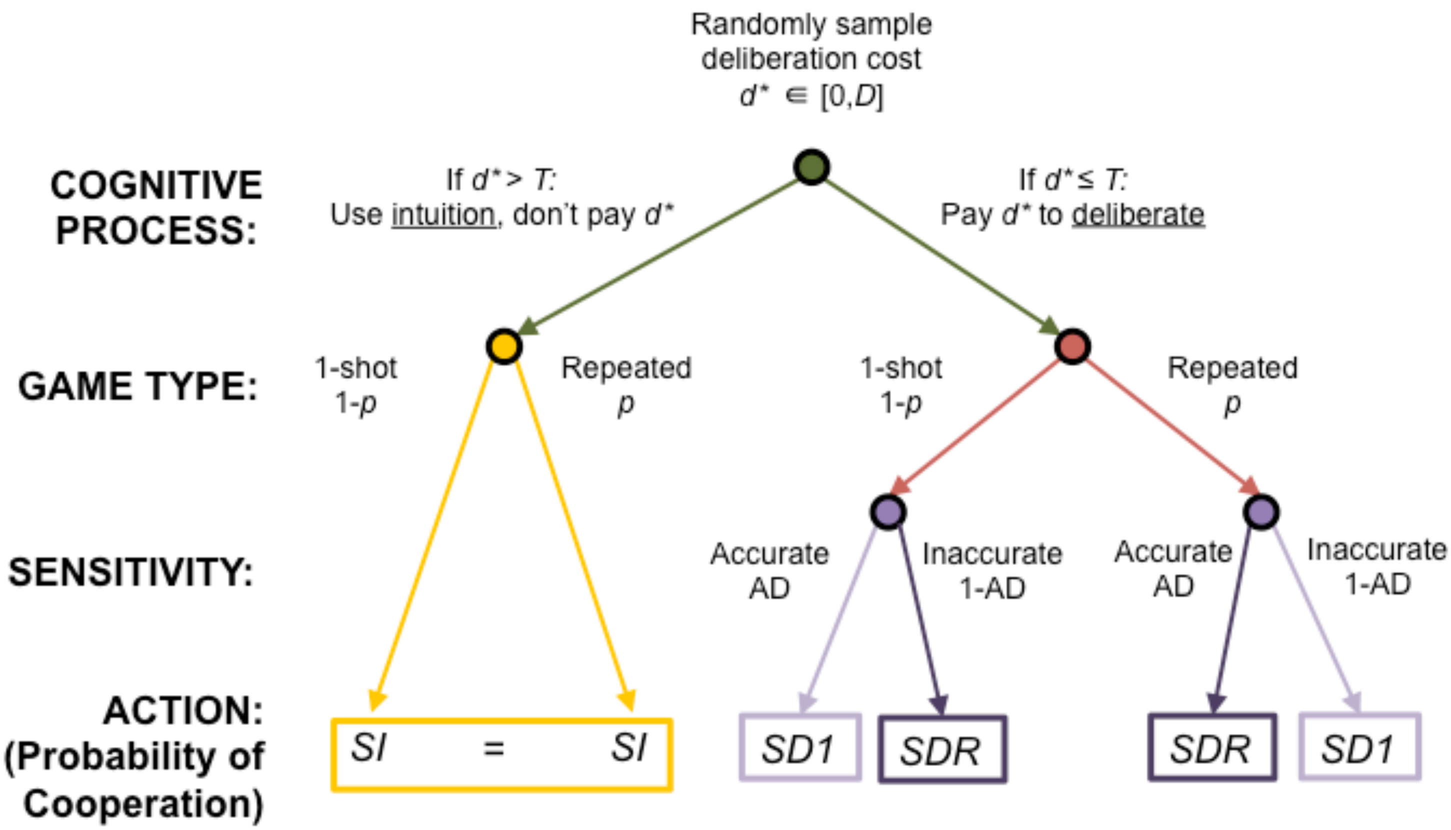

Figure 3. Variable diagram of the model in Study 1. The same four variables, SDI, SD1, SDR, and T are visualized here, along with the sequence of events that takes place in each interaction. First, the agent's cost of deliberation for this interaction d* is sampled uniformly from the interval [0, D]. 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 the agent bases the cooperation decision the generalized intuitive response SI. Since intuition cannot differentiate between game types, the agent plays the cooperative response with probability SI, regardless of whether the game is single-shot (probability 1-p) or repeated (probability p), all identical so far to the original model. On the other hand, if d* ≤ T, deliberation is not too costly, so the agent pays the cost d* and uses deliberation to condition the response on game type. If the agent deliberates accurately (probability AD), then the agent cooperates with probability SD1 if the game is single-shot, and with probability SDR if the game is repeated. However, if the agent deliberates inaccurately (probablity 1-AD), the agent misidentifies the game type, and cooperates with probability SDR if the game is single-shot, and with probability SD1 if the game is repeated.

Results

When imperfect deliberation is introduced to the original model, what effect does it have on the strategies that evolve? I began by varying the accuracy of deliberation (AD) from 1 to 0.5 while p was fixed at 0.5 (once again with B = 4, C = 1, and D = 1).

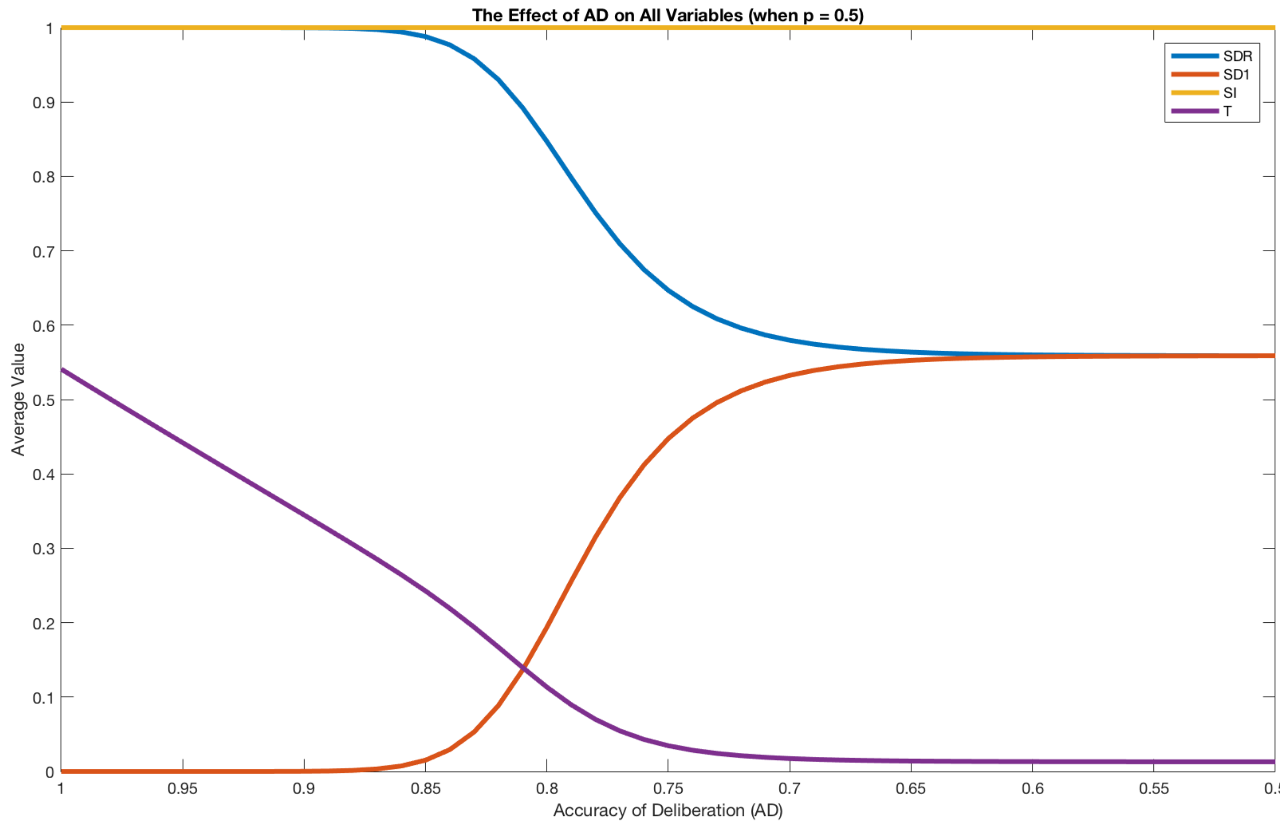

Figure 4. The effect of the accuracy of deliberation on all variables, when p = 0.5. B = 4, C = 1, D = 1. The X axis is the environment variable AD, the accuracy of deliberation. Shown in purple is T, the maximum deliberation cost the agent is willing to pay; in blue is SDR, the agent's probability of cooperating when deliberating and believing they are in a repeated game; in red is SD1, the agent's probability of cooperating when deliberating and believing they are in a single-shot game; in yellow is SI, the agent's probability of cooperating when responding intuitively. The figure was generated using a steady state calculation.

As can be seen in Fig. 4, when deliberation is perfect (AD = 1), agents deliberate slightly more than half of the time (T = 0.5), as in the original model. However, as AD is lowered towards 0.5, the amount of deliberation steadily decreases, hitting floor level near AD = 0.7. Why does decreasing AD reduce the amount of deliberation in the population? As AD is decreased, there becomes a higher chance that the agent will misidentify the game type, and thus play the incorrect response. As a result, the payoff value of deliberation decreases. As the benefit of deliberation decreases, the cost of deliberation, d*, begins to outweigh the benefit, and agents are willing to deliberate less. Thus, as AD decreases, the T value decreases.

The other variables respond accordingly as T declines. Because the probability of being in a repeated game is still 0.5, it is still beneficial for intuition to cooperate, thus SI remains at 1. As AD is decreased and the agents deliberate less, the deliberative responses (SDR and SD1) are used less. As seen in Fig. 2, neutral drift pulls both responses towards 0.5.

The transition that occurs in the graph above as deliberation hits floor level represents a transition from DP-IC to the Intuitive Cooperator (IC), a strategy not discussed in the original model, but one that is in fact present, nonetheless. ICs are agents who have stopped deliberating, but still have a cooperative intuition. In the original model, deliberation hits floor level in the region where the intuitive response to cooperate when $p = 1$. As a result, this represents an extremely small region of IC; however, because deliberation *only* reaches floor once $p = 1$, IC commands such a miniscule portion of the original strategy space that it was not worth discussing in the original findings.

However, it can now be seen that IC is not as insignificant as previously thought. When AD is varied to be less than 1, IC becomes the dominant strategy far before $p = 1$. For example, as can be seen above in Fig. 4, when $p = 0.5$ the dominant strategy transitions to IC once $AD < 0.75$ (representing the threshold at which IC risk dominates DP-IC and deliberation drops to zero). However, this threshold is different with different p values, as can be seen in Fig. 5 below. As p is varied from 0 to 1, the size of AD's effect on T changes. Fig. 5 displays the interaction between p, AD, and T.

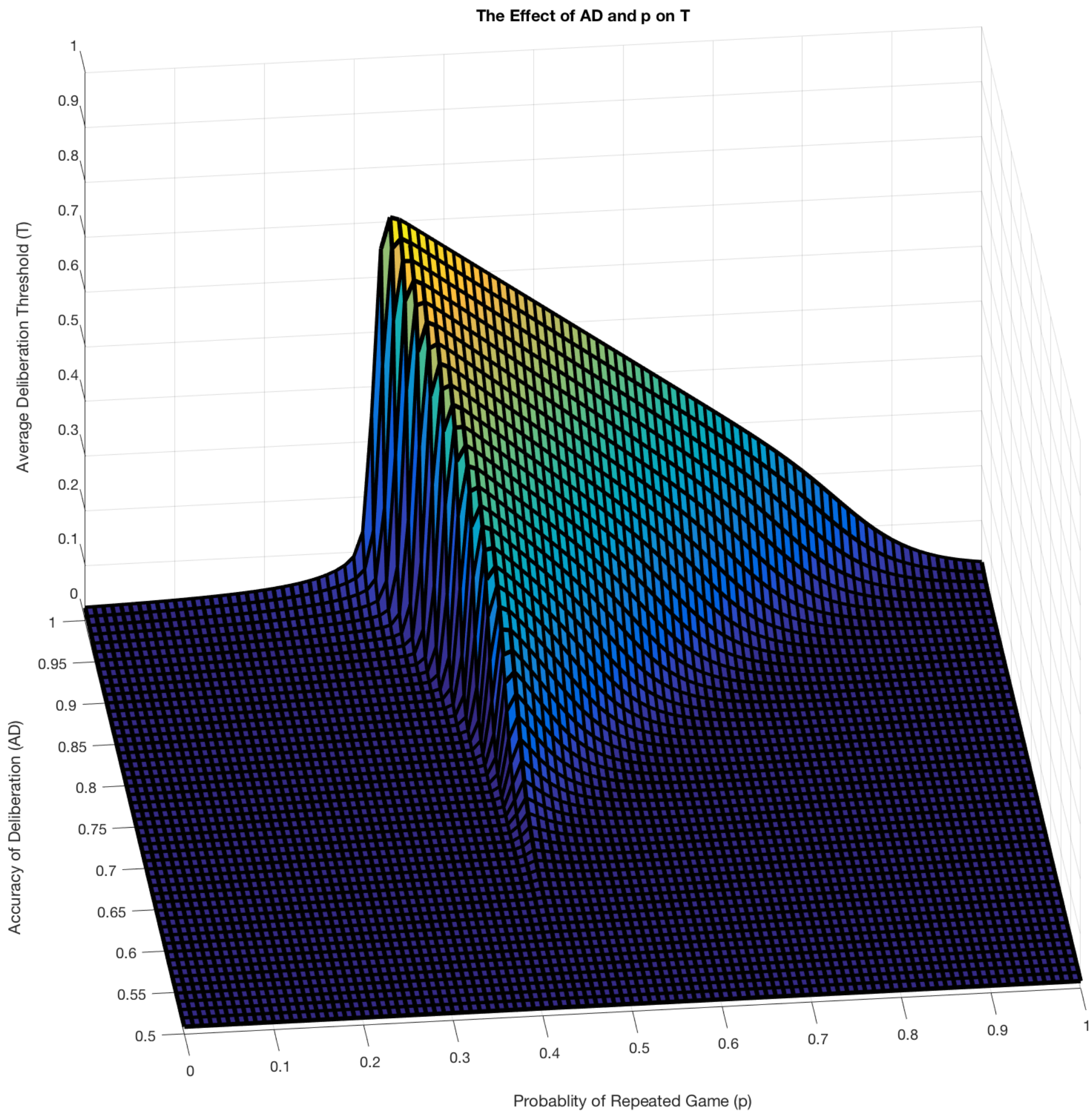

Figure 5. The effect of the accuracy of deliberation and the probability of being in a repeated game on deliberation. B = 4, C = 1, D = 1. The X axis is the environment variable p, the probability of being in a repeated game. The Z axis is the environment variable AD, the accuracy of deliberation. The Y axis is the calculated average value T, the maximum deliberation cost the agent is willing to pay. The higher the "altitude" is at any one [p, AD] coordinate, the greater the level of deliberation is in that environment. The figure was generated using a steady state calculation.

The effects seen here display precisely what one would expect from the combination of the first finding of this study (the effect of varying AD on T) and the finding of Bear & Rand's original model (the effect of varying p on T). Because deliberation was already at floor level in the original model (with AD=1) when p is below 0.3, T cannot decrease any more when deliberation is made less accurate. However, there is an interaction effect as p approaches 1. In the original model, as p approaches 1, deliberation decreases (because the probability of being in a single-shot game becomes lower and lower, thus the chance of deliberation changing the played response decreases), reaching floor level at p = 1. However, as discussed above, IC risk dominates DP-IC at earlier and earlier p values when AD is increased, leading to less and less deliberation in the population.

Because the surface plot above displays T on the y-axis, deliberation takes place anywhere not represented by the flat, navy blue floor. Thus, it is easy to see from this figure where the dual-process agents are. However, this figure fails to distinguish between the IC and ID agents, as neither strategy profile involves the use of deliberation. To better understand where the transitions between dominant strategies take place, we must turn to a strategy space diagram. Fig. 6 displays the dominant strategy in each environment, as p and AD are varied.

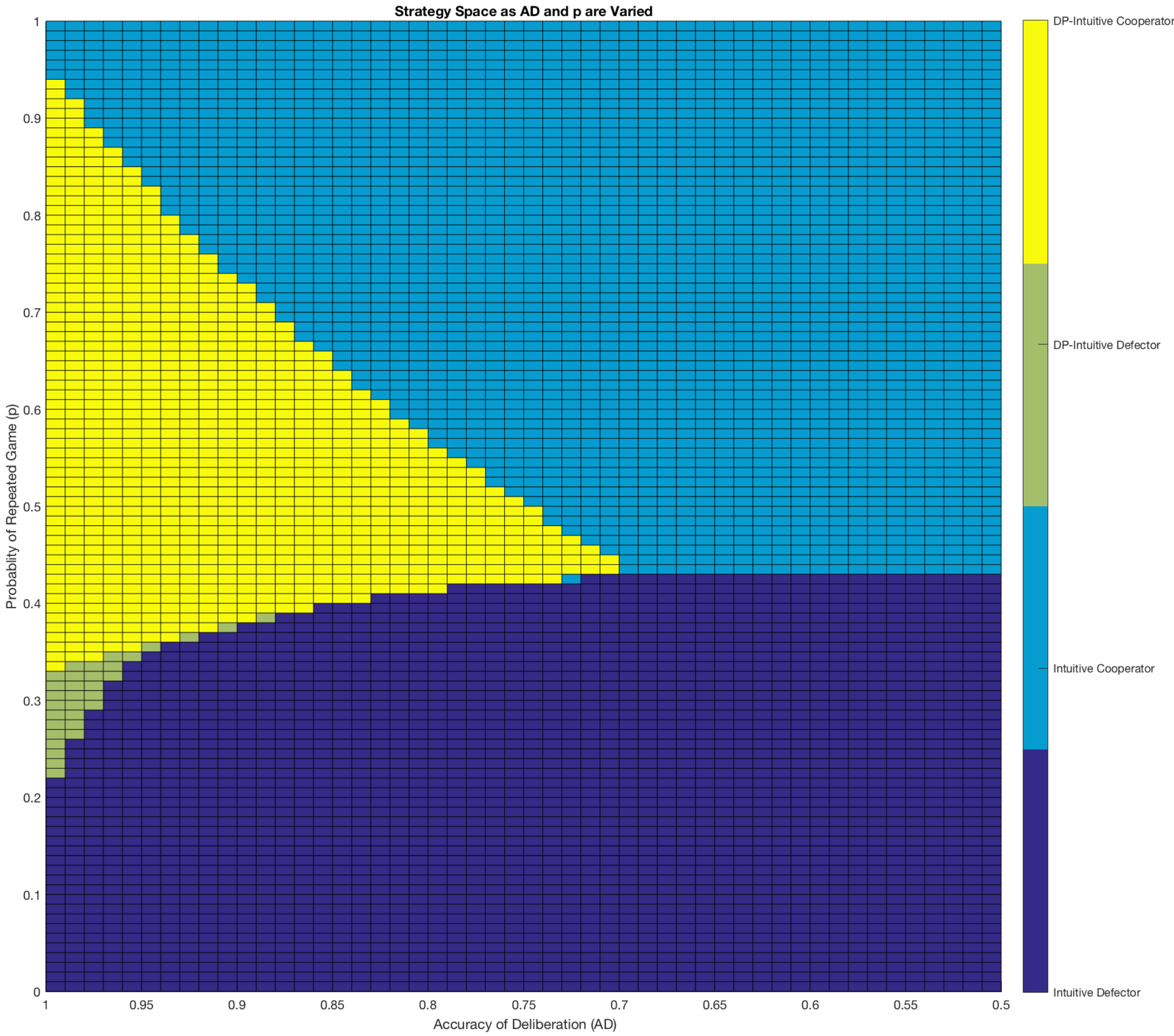

Figure 6. Strategy space diagram as the accuracy of deliberation and probability of being in a repeated game are varied. B= 4, C = 1, D = 1. The X axis is the environment variable AD, the accuracy of deliberation. The Y axis is the environment variable p, the probability of being in a repeated game. The color of each coordinate represents the dominant strategy profile in any one environment, meaning the strategy profile played by more than 50% of the population, determined by looking at the 4 individual variable values in each environment. The light blue region represents the environments dominated by the Intuitive Cooperator strategy. The dark blue region represents the environments dominated by the Intuitive Defector strategy. The yellow region represents the region dominated by the Dual-process Intuitive Cooperator strategy. The slight green region represents the environments dominated by the Dual-process Intuitive Defector strategy. The figure was generated using a steady state calculation.

Fig. 6 displays many of the same effects discussed with the figures above, but instead of displaying these effects in terms of their component variables (T, SDR, SD1, and SI), it displays them in terms of the more macro strategy profiles they make up, revealing the shifts in equilibrium strategies.

The borders between regions of DP-IC, IC, and ID in Fig. 6 represent when one strategy profile risk dominates the other. The yellow region, which represents the region in which DP-IC is the dominant strategy, depicts the same set of environments displayed as any above 0 altitude in Fig. 5, as deliberation only takes place when DP-IC is the dominant strategy. As discussed previously, while there is little representation of IC in the original model (depicted by the left side of Fig. 6), the IC strategy actually represents more than one third of the strategy space once AD is taken into consideration.

On the other hand, the transition between the intuitive response defecting and the intuitive response cooperating (represented by the border between ID and DP-IC/IC) remains approximately constant at 0.3. As in the original model, the region in which the intuition is to defect is far smaller than the region in which the intuition is to cooperate.

In some sense, the relatively greater portion of the strategy space controlled by some form of intuitive cooperator than some form of intuitive defector is related, in a way, to the proclivity for DP-IC to be invaded by IC as deliberation becomes slightly inaccurate. Defecting when you should cooperate is worse than cooperating when you should defect (as set by the ratio of the environment variables, B and C, which determine the payoffs of defecting and cooperating). Evolution is predisposed towards intuitive cooperation. As a result, intuitive cooperation risk dominates intuitive defection at reasonably low p values, and IC risk dominates DP-IC at relatively low amounts of deliberative error. [7]

---

[7] The astute observer will notice the presence of a small region of green, representing DP-ID, which are claimed not to exist by the original model. This is believed to be an error, and is discussed at the end of this paper in the general discussion section.

Study 2: Context-sensitive Intuition

Methods

In the second modification of the model, I introduce the possibility of context-sensitive intuition.[8] To introduce context-sensitive intuitions, the single context insensitive intuitive response (SI) was replaced with two new context-sensitive intuitive responses: a single-shot intuitive response (SI1), and a repeated game intuitive response (SIR). A new environment variable, AI, is introduced to represent the accuracy of intuition. When the agent forgoes deliberation (because d* is greater than T) and plays an intuitive response, the agent will appropriately identify the game type based on the context with some probability, AI. AI was varied as the independent variable in Study 2 along the spectrum from 0.5 to 1. When AI is set to 1, the agent correctly identifies the game type based on the context intuitively 100% of the time, playing SI1 in a single-shot game and SIR in a repeated game (thus obviating any need for deliberation). On the other end of the spectrum, when AI is set to 0.5, the agent fails to ever intuitively identify the game type based on the context and essentially randomly selects between the two intuitive responses. Because there is no higher likelihood of playing the appropriate intuitive response than playing the incorrect response, no selection pressure separates two intuitive responses. As a result, an AI of 0.5 is equivalent to having a single intuitive response (as in the original model). As before with AD, an AI value of 0 represents a perfectly inaccurate response, but is still perfectly informative, so AI is represented on a 0.5 to 1 scale.

[8] For the purpose of isolating the effect of this manipulation, the possibility of imperfect deliberation is once again removed (AD = 1).

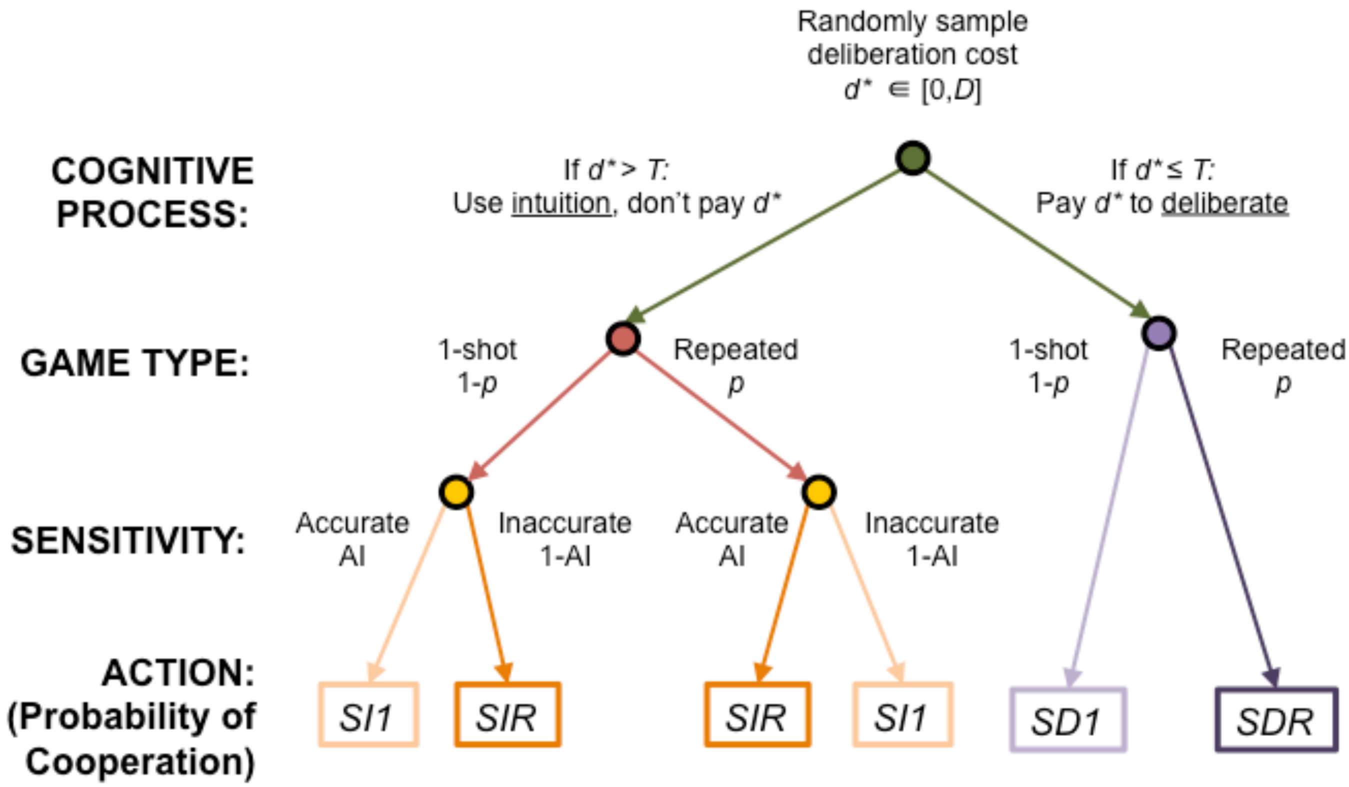

Figure 7. Variable diagram of the model in Study 2. Five variables, SDR, SD1, SIR, SI1 and T are visualized here, along with the sequence of events that takes place in each interaction. (Note the addition of the two context-sensitive intuitions). First, the agent's cost of deliberation for this interaction 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 the agent bases the cooperation decision the intuitive responses SI. Since intuition is now able to partially discriminate between game types, the agent attempts to condition the intuitive response on game type. If the agent accurately discrimiantes between game types (probability AI), the agent cooperates with probability SI1 in a single-shot game (probability 1-p), and with probability SIR in a repeated game (probability p). On the other hand, if the agent fails to accurately discrimiante between game types (probability 1-AI), the agent misidentifies the game type, and cooperates with probablity SIR in a single-shot game, and SI1 in a repeated game. If d* ≤ T, deliberation is not too costly, so the agent pays the cost d* and uses deliberation to condition the response on game type. The agent cooperates with probability SD1 if the game is single-shot, and with probability SDR if the game is repeated.

Results

When the two context-sensitive intuitions are introduced into the original model, what are the resulting strategies? The findings here are slightly more complex than in the case of imperfect deliberation. I began by varying the accuracy of intuition (AI) from 0.5 to 1, fixing the probability of being in a repeated game, p, once again at 0.5 (and once again B = 4, C = 1, and D = 1).

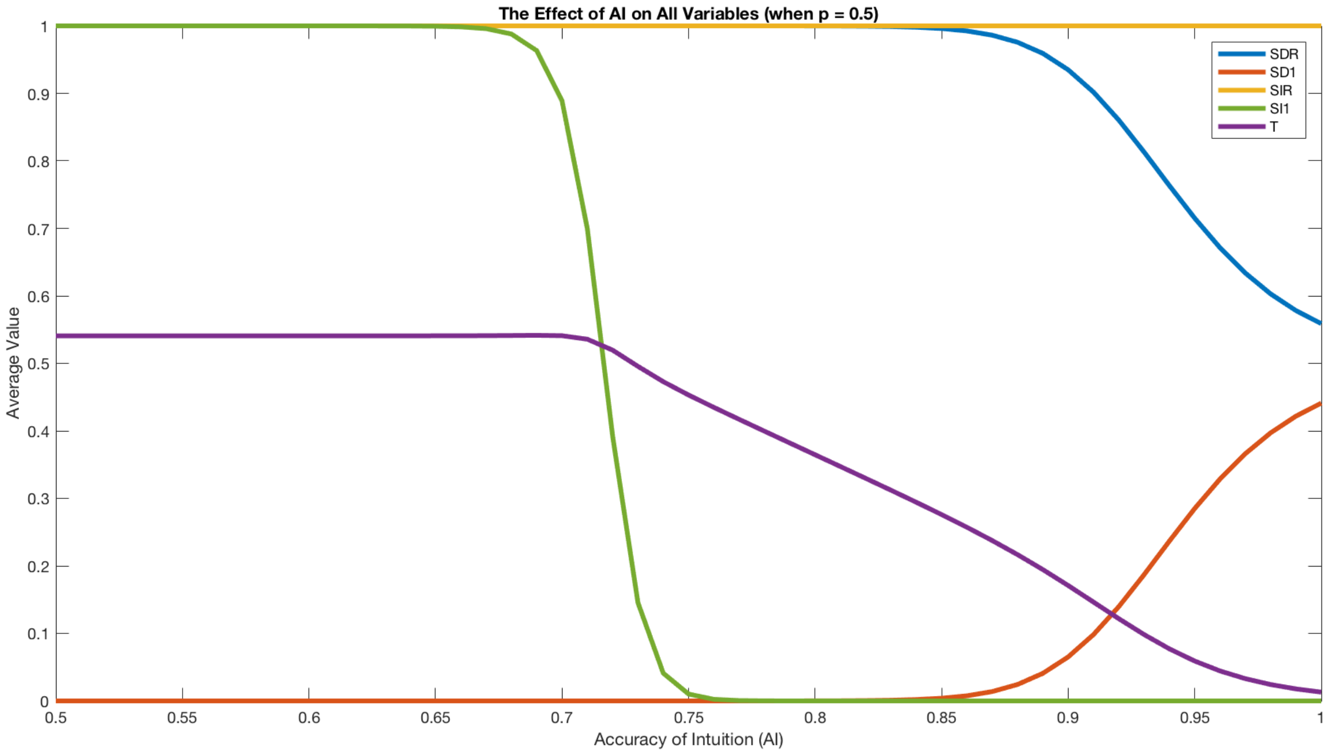

Figure 8. The effect of the accuracy of intuition n all variables, when p = 0.5. B = 4, C = 1, D = 1. The X axis is the environment variable AI, the accuracy of intuition. Shown in purple is T, the maximum deliberation cost the agent is willing to pay; in blue is SDR, the agent's probability of cooperating when deliberating and believing they are in a repeated game; in red is SD1, the agent's probability of cooperating when deliberating and believing they are in a single-shot game; in yellow is SIR, the agent's probability of cooperating when responding intuitively and believing they are in a repeated game; in green is SI1, the agent's probability of cooperating when responding intuitively and believing they are in a single-shot game. The figure was generated using a steady state calculation.

Let us begin by looking at the average deliberation threshold (T), shown by the purple line in Fig. 8. When intuition is insensitive (AI = 0.5), agents deliberate slightly more than half of the time (T = 0.5), as in the original model. As AI is increased, the amount of deliberation remains fixed until approximately AI = 0.7. However, once AI passes 0.7, the amount of deliberation steadily decreases.

The other variables respond accordingly once deliberation begins to decline. Neutral drift pushes the deliberative responses (SDR and SD1) towards 0.5, as they no longer affect fitness. Once deliberation decreases and the intuitive responses are used more frequently, the intuitive responses quickly separate into the appropriate responses of cooperation in a repeated game (SIR) and defection in a single-shot game (SI1). Prior to that point, the intuitive responses both remained constant at 1, representing a single intuitive response of cooperation as seen in the original model when $p = 0.5$.

Why does T decline once AI > 0.7? As AI increases and intuitions become more accurate, agents can discriminate between game types and achieve the same payoff by using context-sensitive intuitions as they would by deliberating. Because agents must pay a cost to deliberate, if they can get the same benefit without paying the cost, they will do so. Once intuitions are sensitive enough to game type, there is no longer a benefit to deliberating. Thus, as AI increases, the T value decreases.

Why is there no decline in T until AI > 0.7? When AI < 0.7 (and greater than 0.5), while intuition is more context-sensitive than in the original model, intuition is still extremely inaccurate, serving only as a very rough tool to discriminate between single-shot and repeated games. Because of this inaccuracy, evolution favors *completely ignoring* the partial context-sensitivity, cooperating regardless of the "perceived" game type, because of the high probability of error. Since the cost of defecting in a repeated game is so high compared to the cost of cooperating in a single-shot game, when there is a risk of misidentifying the game type, it is better to always cooperate and ignore the believed game type, rather than risk accidentally defecting in a repeated game. As a result, the actual intuitive strategy is no different from when AI = 0.5 (and therefore neither is T). It is only once AI is high enough and the context-sensitive intuitions are more accurate that selection begins to favor defecting in a game identified by intuition as single-shot. Once this happens deliberation finally begins to decline, as the comparative benefit of deliberation is reduced.

Once this transition takes place, this represents the transition to a new dominant strategy. Fascinatingly, this dominant strategy is an entirely new strategy favored by evolution not present at all in the original model. These agents use some deliberation, so it is clear that they are some type of dual-processors. However, they are not DP-ICs, with a constant intuition to cooperate occasionally overridden by a deliberative response to defect. Nor is the agent even a DP-ID, with a constant intuition to defect occasionally overridden by a deliberative response to cooperate. Indeed, *there is no constant intuition to override at all* – the agent's intuition discriminates between single-shot and repeated games, as does the agent's deliberation. As a result, the deliberative response does not need to override the intuitive response: the intuitive and the deliberative responses are in harmony for both single-shot and repeated games. In a single-shot game the intuitive response is to defect, as is the deliberative response. In a repeated game the intuitive response is to cooperate, as is the deliberative response. Unlike the DP-IC who only pays for deliberation to override an incorrect intuitive response (allowing the agent to defect in a single-shot game instead of using the intuitive response to cooperate), this new *Harmonious* Dual-processor (HD) only pays to deliberate for the greater accuracy that deliberation provides the agent over the sometimes inaccurate context-sensitive intuitions. The HD represents an entirely new strategy profile not present anywhere in the original model.

The discovery of HD in the new model has some surprising implications. The most surprising effect occurs when p is less than (or equal to) 0.3. In this zone in the original model, ID is the dominant strategy, as deliberation is at floor level and the intuitive response is defect. Increasing AI results in a decrease in deliberation when $p = 0.5$, but as deliberation is already at floor level when $p < 0.3$, it cannot decrease anymore. As a result, one might assume that increasing AI would have no effect when $p < 0.3$.[9] However, this is not in fact the case - when AI is increased from 0.5 to 1 while p = 0.3, there is a large *increase* in deliberation (reaching T = 0.5 at a mere AI = 0.6) before it declines to zero once again.

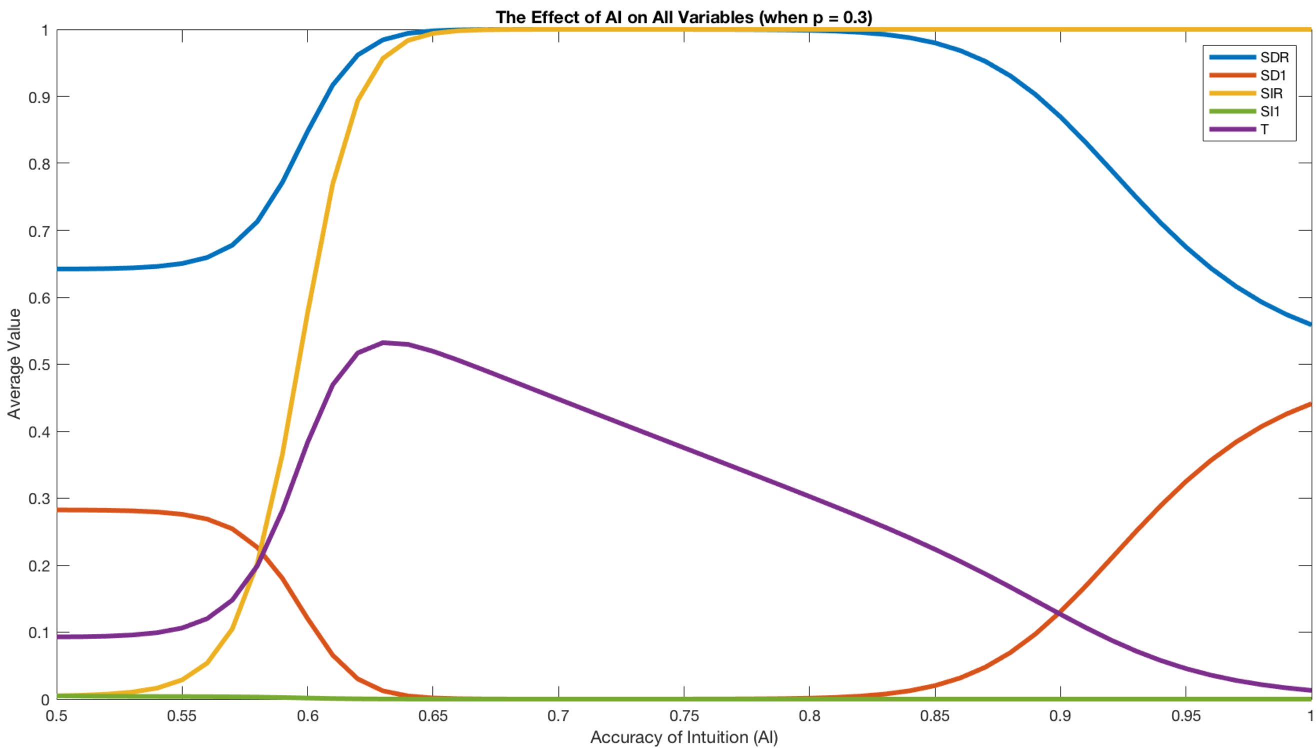

Figure 9. The effect of the accuracy of intuition n all variables, when p = 0.3 B = 4, C = 1, D = 1. The X axis is the environment variable AI, the accuracy of intuition. Shown in purple is T, the maximum deliberation cost the agent is willing to pay; in blue is SDR, the agent's probability of cooperating when deliberating and believing they are in a repeated game; in red is SD1, the agent's probability of cooperating when deliberating and believing they are in a single-shot game; in yellow is SIR, the agent's probability of cooperating when responding intuitively and believing they are in a repeated game; in green is SI1, the agent's probability of cooperating when responding intuitively and believing they are in a single-shot game. The figure was generated using a steady state calculation.

What explains the rise of deliberation in a population that previously was only using intuition, as the accuracy of intuition is increased? The figure above reveals that this decline in deliberation does not represent the transition to a dominant strategy of DP-ID. Instead, it once

[9] In a similar manner to increasing AD when $p < 0.3$ and deliberation is already at floor level

again represents the transition to the HD. The harmonious coordination of SDR with SIR and of SD1 with SI1 can be seen approximately when AI > 0.6.

Why does HD allow for deliberation in a region where there previously was none? Indeed, to stress this once again, this seems rather surprising. One might not expect that making intuition able to discriminate between game types and thereby *more* useful would result in an increase in *deliberation* (and therefore decrease in use of intuition). To understand this effect, we first must recall the finding of Bear & Rand's original research: Dual-process Intuitive Defectors are not favored by evolution because it is only worthwhile to switch to cooperation when the other agent also deliberates (and thus also switches to cooperate). Therefore, agents maximize their payoff by deliberating less than their partner, ultimately resulting in zero deliberation (explained in detail in the previous section on the original model).

However, as AI is increased and intuitions become context-sensitive, the SIR response allows agents to cooperate in repeated games without deliberating. So, if an agent deliberates and identifies a repeated game, there is now a higher chance that their partner will have chosen to cooperate as well, due to the cooperation generated through SIR. This means that even if their partner never *deliberates*, there is still some potential gain from deliberating and switching to cooperation for the original agent. As a result, the optimal strategy has T > 0. As seen in Fig. 8 when p = 0.5 and AI > 0.7, agents are once again sometimes willing to pay for the increased accuracy that deliberation can provide. Thus, instead of the Intuitive Discriminator being the dominant strategy, deliberation is introduced into the population once more, and the Harmonious Dual-processor is the dominant strategy.

On the other hand, as AI is increased further and approaches 1, the accuracy of SIR approaches perfection, and agents no longer get higher accuracy by deliberating. SIR serves as a cheaper method of achieving the same response as SDR, and deliberation once again declines (as seen before in Fig. 8 when p = 0.5 and AI > 0.7). Eventually deliberation hits floor level, and there is a brief period in which the Intuitive Discriminator is technically the dominant strategy. However, much like the Intuitive Cooperator was not present in the original model, the Intuitive Discriminator does not play a large role in this model, as in both cases deliberation only reaches floor level for an extremely small portion of the strategy space. The interaction between p, AI, and T can be observed in the three dimensional surface plot below.

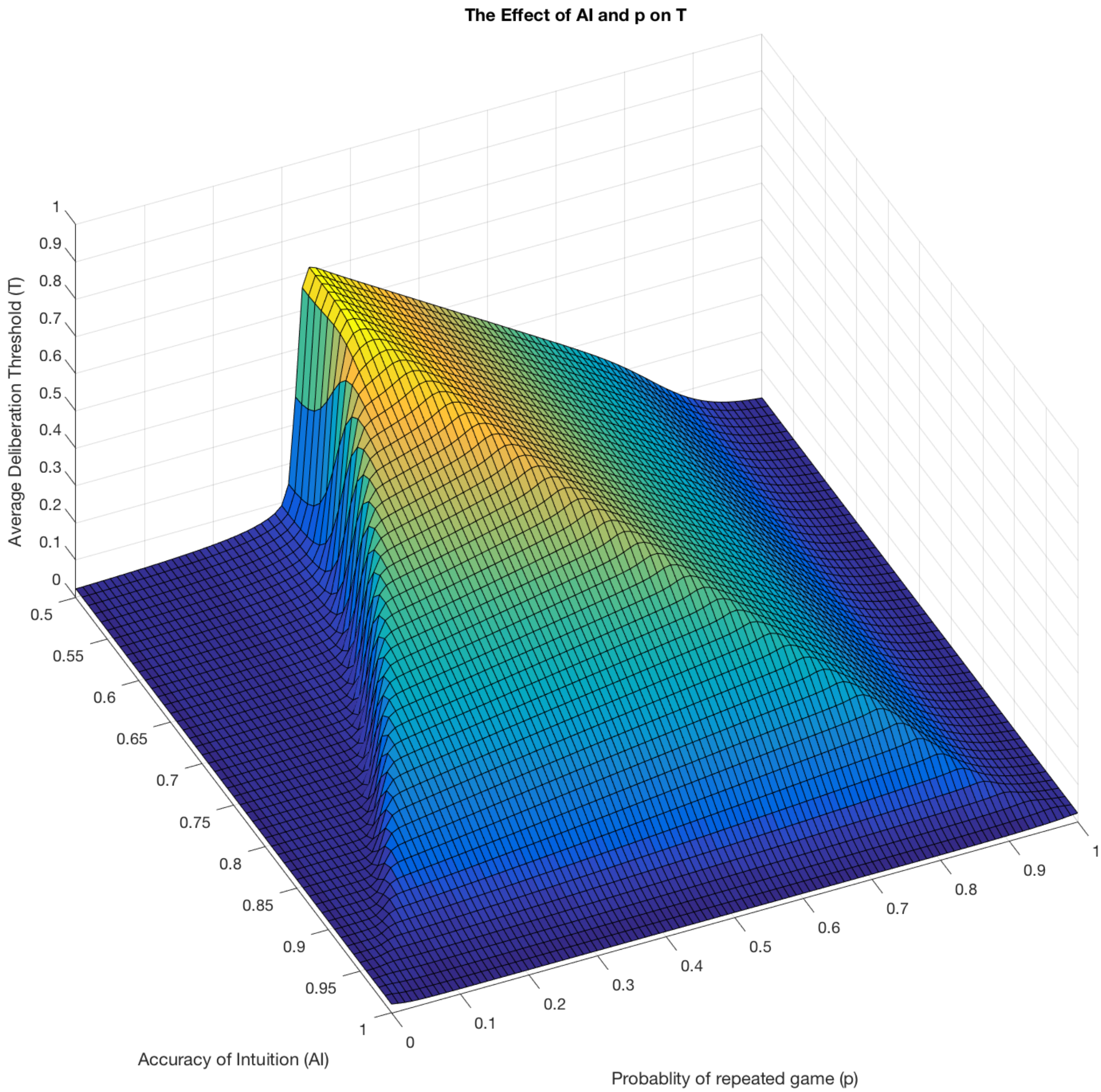

Figure 10. The effect of the accuracy of intuition and the probability of being in a repeated game on deliberation. B = 4, C = 1, D = 1. The X axis is the environment variable p, the probability of being in a repeated game. The Z axis is the environment variable AI, the accuracy of intuition. The Y axis is the calculated average value for T, the maximum deliberation cost the agent is willing to pay. The higher the "altitude" is at any one [p, AI] coordinate, the greater the level of deliberation is in that environment. The figure was generated using a steady state calculation.

Once again (as in Fig. 5), any altitude above floor level represents the existence of deliberation, and therefore a dual-process agent. One face of the figure represents the DP-IC, the other represents the HD, the border representing the point at which one risk dominates the other. The relationship between HD and the other strategies favored by evolution can better be understood by examining the strategy space diagram below.

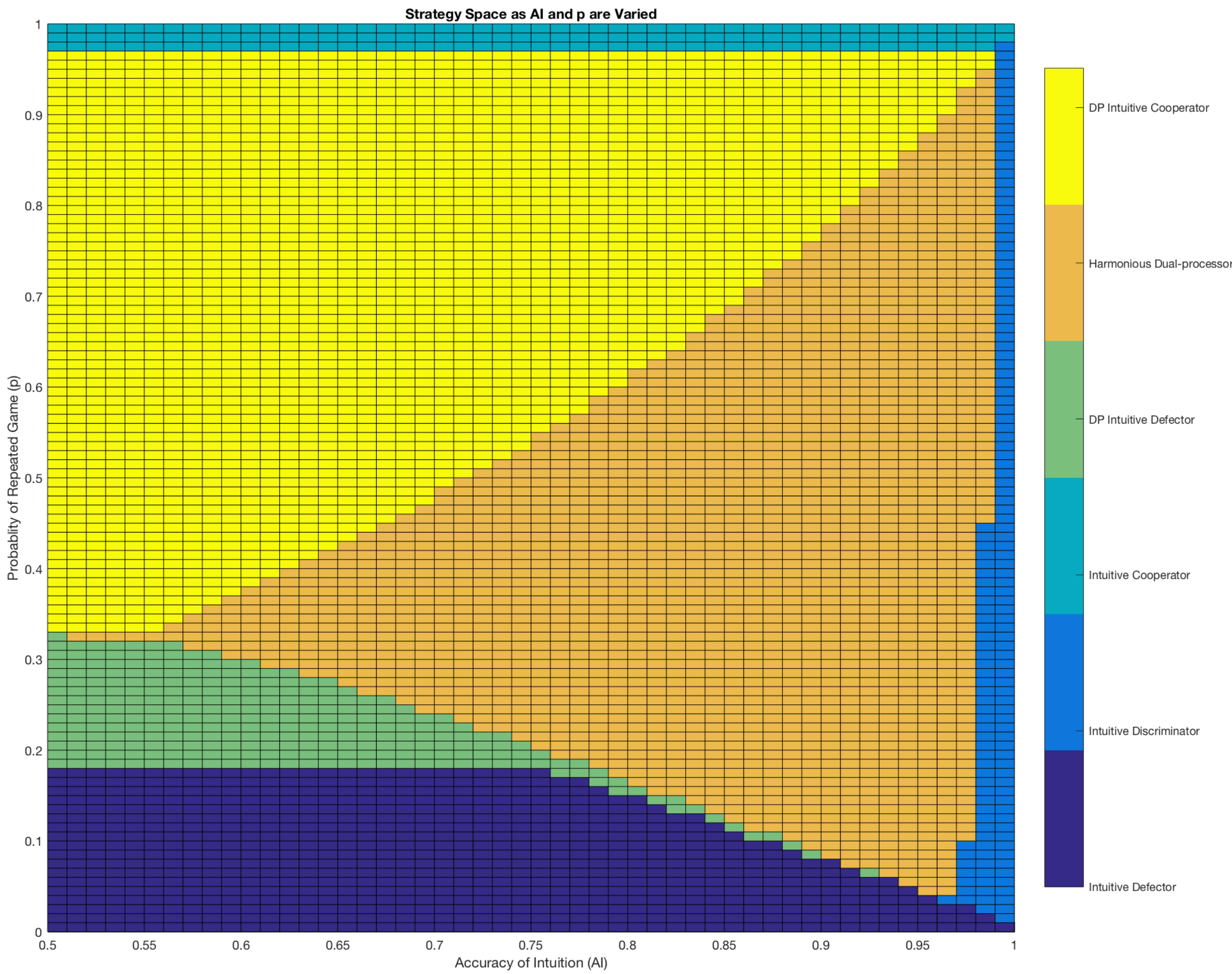

Figure 11. Strategy space diagram as the accuracy of intuition and probability of being in a repeated game are varied. B= 4, C = 1, D = 1. The X axis is the environment variable AI, the accuracy of intuition. The color of each coordinate represents the dominant strategy profile in any one environment, meaning the strategy profile played by more than 50% of the population, determined by looking at the 5 individual variable values in each environment. The teal region represents the environments dominated by the Intuitive Cooperator strategy. The dark blue region represents the environments dominated by the Intuitive Defector strategy. The yellow region represents the region dominated by the Dual-process Intuitive Cooperator strategy. The light blue region represents the environments dominated by the Intuitive Discriminator strategy. The orange region represents the environments dominated by the new Harmonious Dual-processor strategy. The slight green region represents the environments dominated by the Dual-process Intuitive Defector strategy. The figure was generated using a steady state calculation.

In Fig. 11, the full extent of the Harmonious Dual-processor can truly be seen, as a large portion of the strategy space is dominated by this strategy. The Intuitive Discriminator can also be seen once AI almost reaches exactly 1. In this region, the context-sensitive intuitions are so accurate that there is no need for deliberation in the population – the intuitions perfectly discriminate between game types.

General Discussion

By introducing imperfect deliberation and context-sensitive intuitions into a dual-process game-theoretic model of the evolution cooperation, I was able to test the ecological validity of the original Bear & Rand model. The findings of the current research demonstrate that the original findings are reasonably robust, maintaining their validity even under a decent amount of modification to the model. Furthermore, the two new models allow for the study of the evolution of cooperation in dual-process agents in a larger variety of contexts than allowed for by the original model.

With respect to imperfect deliberation, the finding was that as the accuracy of deliberation decreases, the amount of deliberation in the population decreases as well. However, even with some imperfect deliberation, the two main findings of the original research still hold up to some point. That is to say, Intuitive Defector still remains the dominant strategy when there is a low probability of being in a repeated game and DP-IC remains the dominant strategy when there is a higher probability of being in a repeated game (up until significant levels of inaccuracy or very high probabilities of being in a repeated game).

The implication is that relaxing the perfection of deliberation in the original Bear & Rand model does not lead to invalidation of the results, unless deliberation is quite inaccurate. While it is likely that deliberation is not *perfect* in the real world, it is also unlikely that deliberation is extremely inaccurate. If deliberation is tolerably accurate, it can still be asserted that ID and DP-IC will remain the dominant strategies in their respective domains.

However, the current research also reveals that the less accurate deliberation is, the lower the p value needs to be for IC to become the dominant strategy. Although when the probability of being in a repeated game is moderately low there is typically still deliberation (provided AD is not extremely low), deliberation drops to zero when there is a higher probability of being in a repeated game. This effect is seen in the original model as well – at very high probabilities of a repeated game, the benefit of deliberation decreases (as there is little chance of deliberation altering the strategy, which only would occur in a single-shot game) and is surpassed by the benefit of always cooperating. However, with the additional inaccuracy of deliberation introduced by this model, the effect takes place even sooner, leading deliberation to reach floor level sometimes at a slightly higher probability of being in a repeated game than the threshold at which DP-IC risk dominates ID. Thus, this model shows that IC is the dominant strategy in a larger number of environments than predicted by the original model.

By introducing context-sensitive intuitions into the original Bear & Rand model, I was able to test the ecological validity of the original model in a second dimension. Here the findings were more interesting. In one sense, the robustness of the original model was upheld. The observation of a significantly sized plateau in the amount of deliberation as AI is increased demonstrates that the original findings regarding DP-ICs still hold, even with a significant increase in the accuracy of intuitions. Agents ignore the intuitive sensitivity to single-shot vs. repeated games, up to a point.

On the other hand, the discovery of an entirely new strategy favored by evolution, the Harmonious Dual-processor (once AI is increased sufficiently and the intuitive responses separate), is quite fascinating. This strategy controls a very large portion of the strategy space, despite not being present in the environments explored by the original model at all.

This is, perhaps, the most significant discovery of these studies. The discovery of HD is not merely another manifestation of the same dual-processing, as perhaps the discovery of DP-ID would have been. The discovery of HD shows that System 1 and System 2 do not need to be in conflict with each other, as most models of dual-process theory present them to be. Instead of the standard model where System 2 overrides the System 1 intuition, providing self-control, an HD uses System 1 and System 2 to advance the exact same goal – a response appropriately conditioned on the game type.

The discovery of HD also had some fascinating implications for the ID. HD allows for the evolution of deliberation in environments where there previously was none, suggesting that the number of possible environments in which ID evolves to be the dominant strategy is much smaller than predicted by the original model. Instead, when a small amount of sensitivity is introduced into the intuitions, ID is overrun by HD, agents who pay to deliberate for higher accuracy.

What does the current research suggest about actual human behavior? It suggests that while in the real world we may very well have somewhat context-sensitive intuitions and occasionally imperfect deliberation, the majority of the time this does not affect our decision to deliberate or intuit, cooperate or defect. Because the repercussions of mistakenly defecting are typically far greater than the repercussions of mistakenly cooperating (Delton et al., 2014), we have primarily evolved to use our intuition to cooperate most of the time (even when we have somewhat context-sensitive intuitions), and occasionally we choose to deliberate and defect if in an interaction without reciprocal consequences (so long as our ability to deliberate is reasonably intact).

However, the two models developed by the current research, that of imperfect deliberation and context-sensitive intuitions, can also be used to understand a number of scenarios that are not explained by the original model. Because we may have evolved to use different strategy profiles in different meta-contexts, we may be Dual-process Cooperators a majority of the time, but sometimes act as Harmonious Dual-processors or Intuitive Cooperators (strategies that the two new models demonstrate dominate in a larger number of environments than predicted by the original model).

For instance, consider the practice of stereotyping – creating generalizable impressions about a category of people that can be inferred (correctly or not) from a few indicative characteristics. Stereotyping provides context for otherwise unknown situations, allowing agents to make fast, System-1, context-sensitive intuitive responses. Judgments about whether to help a stranger may be influenced by our stereotypes about them (System 1 intuitions). However, we also sometimes stop and analyze our stereotypes, considering whether the context upon which we are basing our decision is actually providing relevant information. This is making use of System 2 deliberation to provide a more accurate analysis of the situation than the somewhat context-sensitive System 1 intuition can. Thus, we can see that stereotyping provides an excellent example of an environment in which we act as a Harmonious Dual-processor, sometimes conditioning our response intuitively and sometimes choosing to deliberate for greater accuracy.

The models developed by the current studies also present several testable predictions. One prediction is that when deliberation is imperfect and subjects are placed in the context of a single-shot game, they should rely even more heavily on intuition than previously found in single-shot games with perfect deliberation, resulting in a higher frequency of cooperation despite the single-shot game type. One way of testing this would be to provide subjects with intentionally ambiguous information regarding the game type, thereby reducing the accuracy of deliberation.

A second prediction made by the current research is that even when subjects have intuitive beliefs regarding the game type (context-sensitive intuitions), there should still be a high frequency of intuitive cooperation in single-shot games (as the intuitions do not separate unless those intuitive beliefs are very accurate at discriminating between game types). One way to test this prediction would be to train subjects to intuitively/subconsciously interpret non-obvious context clues regarding the game type (to create context-sensitive intuitions), and then test their intuitive responses (using time pressure) when in a single-shot game. If the prediction of this model is correct, as long as the context clues are not extremely informative of game type, subjects should still intuitively cooperate when in a single-shot game, ignoring the additional information.

Even though two of the simplifications of the original model have been removed through these studies, the models presented here still include a number of constraints and simplifications. One such simplification is that the current research assumed that the cost, d*, is sampled from a uniform distribution on the interval 0 to D. However, a better model might include a more realistic distribution of costs.

Additionally, the studies only allow for two payoff structures: single-shot games, and games with reciprocal consequences (such as repeated games). However, it could be interesting to consider how agents would respond to a distribution of increasing reciprocal consequences for defecting, for example modeling the possibility of longer and longer repeated games. Furthermore, if such a model were implemented it could be interesting to also allow for an asymmetric ability to misidentify the game type. Study 1 assumes that agents are equally likely to misidentify repeated games and single-shot games as AD is increased. However, a more realistic model with many game lengths/payoff structures might make it easy for agents to accurately identify single-shot games, but increasingly difficult for agents to determine the length of longer and longer games.

Finally, an extension of the model could include a way for agents to predict (with some probability of accuracy) whether their partner intends to cooperate or defect, making it is easier to coordinate. One method of implementation would be to make deliberation observable (as it is often time consuming) such that an agent could modify their response conditional upon the cognitive behavior of their partner. Alternatively, agents could develop visible reputations based on their behavior in previous interactions, allowing agents to guess what their partner is likely to do in the current game.

Future research should also investigate the observed population of Dual-process Intuitive Defectors that the original model claims does not exist, but in fact can be observed in an extremely small portion of the environments in both the current studies and the original Bear & Rand model. Such a population can be seen in the original findings at low probabilities of being in a repeated game (where the intuitive response is to defect). Seeing as deliberation is not actually perfectly at floor level, a population of agents must exist that very occasionally uses deliberation even when their intuition is to defect. This population of agents persists through the manipulations done in the current research, and can be seen as the green region in both strategy space figures.

In sum, I introduced two complications to the dual-process game-theoretic model of the evolution of cooperation developed by Bear & Rand, to increase the ecological validity of the original model and to allow for its application to new contexts. I introduced the possibility of imperfect deliberation, in which the agent deliberates, misidentifies the game type, and plays the wrong response, and I introduced the possibility of context-sensitive intuition, in which the agent conditions their intuitive response on the game type without deliberating, thereby gaining the benefit of deliberation without the cost. In doing so, I demonstrated that the original model is quite robust: even with some amount of imperfect deliberation and with a substantial amount of context-sensitive intuition, the original model still holds to a strong degree. However, I also demonstrated that context-sensitive intuitions can lead to the development of deliberation in a region where no deliberation existed previously. Finally, in the most surprising discovery of all, the introduction of context-sensitive intuitions revealed a new strategy favored by evolution, the Harmonious Dual-processor.

References

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390-1396.

Bear, A., & Rand, D. G. (2016). Intuition, deliberation, and the evolution of cooperation. *Proceedings of the National Academy of Sciences, 113*(4), 936–941.

Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences*, *108*(32), 13335-13340.

Duffy, S., & Smith, J. (2014). Cognitive load in the multi-player prisoner's dilemma game: Are there brains in games?. *Journal of Behavioral and Experimental Economics*, *51*, 47-56.

Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol.*, *59*, 255-278.

Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford University Press, USA.

Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *American psychologist*, *58*(9), 697.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Moran, P. A. P. (1958). Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society, 54*(1), 60-71.

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, *314* (5805), 1560-1563.

Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*,*437*(7063), 1291-1298.

Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, *489* (7416), 427-430.

Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in cognitive sciences*, *17*(8), 413-425.

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

Tomlin, D., Rand, D. G., Ludvig, E. A., & Cohen, J. D. (2015). The evolution and devolution of cognitive control: The costs of deliberation in a competitive world. *Scientific reports*, *5*.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly review of biology, 46*(1), 35-57.

Van Veelen, M., García, J., Rand, D. G., & Nowak, M. A. (2012). Direct reciprocity in structured populations. *Proceedings of the National Academy of Sciences*, *109*(25), 9929-9934.