A Resource-Rational Mechanistic Approach to One-shot Non-cooperative Games: The Case of Prisoner's Dilemma

Ardavan S. Nobandegani[1,3]**, Kevin da Silva Castanheira**[3]**, Thomas R. Shultz**[2,3]**, & A. Ross Otto**[3]
{ardavan.salehinobandegani, kevin.dasilvacastanheira}@mail.mcgill.ca
{thomas.shultz, ross.otto}@mcgill.ca
[1]Department of Electrical & Computer Engineering, McGill University
[2]School of Computer Science, McGill University
[3]Department of Psychology, McGill University

Abstract

The concept of Nash equilibrium has played a profound role in economics, and is widely accepted as a normative stance for how people should choose their strategies in competitive environments. However, extensive empirical evidence shows that people often systematically deviate from Nash equilibrium. In this work, we present the first resource-rational mechanistic approach to one-shot, non-cooperative games (ONG), showing that a variant of normative expected-utility maximization acknowledging cognitive limitations can account for important deviations from the prescriptions of Nash equilibrium in ONGs. Concretely, we show that Nobandegani et al.'s (2018) metacognitively-rational model, *sample-based expected utility*, can account for purportedly irrational cooperation rates observed in one-shot, non-cooperative Prisoner's Dilemma, and can accurately explain how cooperation rate varies depending on the parameterization of the game. Additionally, our work provides a resource-rational explanation of why people with higher general intelligence tend to cooperate less in OPDs, and serves as the first (Bayesian) rational, process-level explanation of a well-known violation of the law of total probability in OPDs, documented by Shafir and Tversky (1992), which has resisted explanation by a model governed by classical probability theory for nearly three decades. Surprisingly, our work demonstrates that cooperation can arise from purely selfish, expected-utility maximization subject to cognitive limitations.

Keywords: One-shot non-cooperative games; Nash equilibrium; resource-rational process models; expected utility theory; behavioral game theory; Prisoner's Dilemma; cooperation

1 Introduction

In his seminal work, Nash (1950) introduced a foundational concept of equilibrium, now called "Nash equilibrium," and mathematically proved that any one-shot, non-cooperative, n-player game enjoys (at least) one such equilibrium. In simple terms, Nash equilibrium (NE) is a set of strategies, one for each of the n players of the game, which has the desirable property that each player's strategy is her best response to the strategies adopted by the $n-1$ other players.

Importantly, NE satisfies a number of notable conditions which make it appealing from a normative standpoint. For example, NE passes the key announcement test (Holt & Roth, 2004): If all players publicly announce their strategies, no player would want to reconsider. Furthermore, when the goal is to advise players of a game about which strategies to follow, NE stands out as a rational choice: Any advice that is not an NE would have the unsettling property that there would always be some player(s) who would be better off by deviating from what they are advised (Holt & Roth, 2004). Finally, NE is a self-reinforcing agreement (Holt & Roth, 2004): Once reached by the players, NE does not need any external means of enforcement to endure.

Despite its firm rational grounds, NE has repeatedly failed to provide a descriptively adequate account of human behavior in a variety of important game-theoretic settings (e.g., Mailath, 1998; Goeree & Holt, 2001). By now, extensive empirical evidence shows that people often systematically deviate from Nash equilibrium, thus calling for alternative accounts (e.g., Fehr & Gächter, 2000; Keser & van Winden, 2000; Brandts & Schram, 2001). A prominent example of such violations of NE is the robust empirical finding that people typically cooperate in 2-player, one-shot, non-cooperative, Prisoner's Dilemma (2ONPD) games. Not only does NE prescribe against cooperation in 2ONPD (more precisely, every 2ONPD has only a single NE, and that is for both players to defect), but, more importantly, cooperation is not even *rationalizable* in 2ONPD (Bernheim, 1984; Pearce, 1984) because cooperation is not a best response to any strategy adopted by the other player.

From a purely computational perspective, people's apparent failure to follow the prescriptions of NE is not surprising: Recent theoretical work in computational complexity formally showed that evaluating NE is computationally intractable in general (Daskalakis et al., 2009), and, hence, is generally beyond the capacity of a cognitive system with limited computational power and resources (Simon, 1957).

In this work, we present the first resource-rational mechanistic approach to one-shot, non-cooperative games (ONGs), investigating the extent to which violations of NE could be seen as an optimal response subject to computational and cognitive limitations (Griffiths, Lieder, & Goodman, 2015; Nobandegani, 2017). Concretely, we ask whether these violations can be seen as an optimal behavior with the mind acting as a cognitive miser. To do this, we begin by presenting a general framework allowing us to conceptualize any ONG as a set of risky gambles, thereby reducing the problem of strategy selection in ONGs to a problem of choosing between a set of risky gambles.

To show the efficacy of our framework, we investigate the robust, yet puzzling, experimental finding that people typically cooperate in 2ONPD (e.g., Fehr & Gächter, 2000; Keser & van Winden, 2000; Brandts & Schram, 2001). As we demonstrate, Nobandegani, da Silva Castanheira, Otto, and Shultz's (2018) metacognitively-rational model, *sample-based expected utility* (SbEU), not only can provide a resource-rational mechanistic explanation of cooperation behavior in 2ONPD, but also can provide a remarkably accurate quantitative account of how cooperation rate varies depending on the parameterization of 2ONPD (i.e., specific payoffs of the game).

Our paper is organized as follows. After providing a brief overview of SbEU, we present a general framework permitting us to reduce the problem of strategy selection in ONGs to the problem of decision-making under risk. We then turn to modeling cooperation in 2ONPD. Finally, we conclude by discussing the implications of our work for the debate on human rationality.

2 Sample-based Expected Utility Model

Extending the decision-making model of Lieder, Griffiths, and Hsu (2018) to the realm of metacognition, SbEU is a metacognitively-rational process model of risky choice, positing that an agent rationally adapts their strategies depending on the amount of time available for decision-making (Nobandegani et al., 2018). Concretely, SbEU assumes that an agent estimates expected utility

$$\mathbb{E}[u(o)] = \int p(o)u(o)do, \tag{1}$$

using self-normalized importance sampling (Hammersley & Handscomb, 1964; Geweke, 1989), with its importance distribution q^* aiming to optimally minimize mean-squared error (MSE):

$$\hat{E} = \frac{1}{\sum_{j=1}^{s} w_j} \sum_{i=1}^{s} w_i u(o_i), \quad \forall i:\ o_i \sim q^*,\ w_i = \frac{p(o_i)}{q^*(o_i)}, \tag{2}$$

$$q^*(o) \propto p(o)|u(o)| \sqrt{\frac{1+|u(o)|\sqrt{s}}{|u(o)|\sqrt{s}}}. \tag{3}$$

MSE is a standard normative measure of the quality of an estimator, and is widely adopted in machine learning and mathematical statistics (Poor, 2013). In Eqs. (1-3), o denotes an outcome of a risky gamble, $p(o)$ the objective probability of outcome o, $u(o)$ the subjective utility of outcome o, $\hat{E}$ the importance-sampling estimate of expected utility given in Eq. (1), q^* the importance-sampling distribution, o_i an outcome randomly sampled from q^*, and s the number of samples drawn from q^*.

SbEU posits that, when choosing between a pair of risky gambles A, B, people make their choice depending on whether the expected value of the utility difference $\Delta u(o)$ is negative or positive (w.p. stands for "with probability"):

$$A = \begin{cases} o_A & \text{w.p. } P_A \\ 0 & \text{w.p. } 1-P_A \end{cases} \tag{4}$$

$$B = \begin{cases} o_B & \text{w.p. } P_B \\ 0 & \text{w.p. } 1-P_B \end{cases} \tag{5}$$

$$\Delta u(o) = \begin{cases} u(o_A)-u(o_B) & \text{w.p. } P_A P_B \\ u(o_A)-u(0) & \text{w.p. } P_A(1-P_B) \\ u(0)-u(o_B) & \text{w.p. } (1-P_A)P_B \\ 0 & \text{w.p. } (1-P_A)(1-P_B) \end{cases} \tag{6}$$

In Eq. (6), $u(\cdot)$ denotes the subjective utility function of a decision-maker. Following Nobandegani et al. (2018), and consistent with prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992), in this paper we assume a standard S-shaped utility function $u(x)$ given by:

$$u(x) = \begin{cases} x^{0.85} & \text{if } x \geq 0, \\ -|x|^{0.95} & \text{if } x < 0. \end{cases} \tag{7}$$

Nobandegani et al. (2018) recently revealed that SbEU provides an account of the availability bias, the tendency to overestimate the probability of events that easily come to mind (Tversky & Kahneman, 1973), and can accurately simulate the well-known fourfold pattern of risk preferences in outcome probability (Tversky & Kahneman, 1992) and in outcome magnitude (Markovitz, 1952; Scholten & Read, 2014). Notably, SbEU is the first rational process model to score near-perfectly in optimality, economical use of limited cognitive resources, and robustness, all at the same time (see Nobandegani et al., 2018; Nobandegani et al., 2019a).

3 From One-shot, Non-cooperative Games to Multi-alternative Risky Choice

In this section, we present a general framework allowing us to conceptualize any ONG as a set of risky gambles $\mathcal{S}$. Importantly, this framework permits us to reduce the problem of strategy selection in ONGs to the problem of risky decision-making. By re-framing the problem this way, the strategy selected by an agent in an ONG corresponds to the risky gamble that the agent would choose among the set of available gambles $\mathcal{S}$.[1]

Without loss of generality, and for ease of exposition, we consider the case of a 2-player, one-shot, non-cooperative game (2ONG) here. Extending the results to the general case of an n-player, one-shot, non-cooperative game is straightforward.

Consider a generic 2ONG whose payoff matrix is given in Fig. 1. The game has two players: Player 1 (Row Player) and Player 2 (Column Player). Player 1 has two pure strategies to choose between: the strategy corresponding to choosing the top row (Top Strategy) and the strategy corresponding to choosing the bottom row (Bottom Strategy). Similarly, Player 2 has two pure strategies to choose from: the strategy corresponding to choosing the left column (Left Strategy) and the strategy corresponding to choosing the right column (Right Strategy). From the perspective of Player 1, Player 2 selects the Left Strategy with probability P_l, and the

[1] We should note that our framework naturally handles "mixed strategies" wherein the agent probabilistically chooses among the set of possible "pure strategies." The validity of this claim follows from the key understanding that the choice between the set of available gambles $\mathcal{S}$ would be also made probabilistically.

		Player 2 (Column Player)	
		Left	Right
Player 1 (Row Player)	Top	a, x	b, v
	Bottom	c, y	d, w

Figure 1: Payoff matrix for a generic 2-player, one-shot, non-cooperative game (2ONG). For example, if Player 1 (Row Player) selects the Top Strategy and Player 2 (Column Player) selects the Left Strategy, Player 1 and Player 2 receive payoffs a and x, respectively.

Right Strategy with probability $P_r = 1 - P_l$. Likewise, from the perspective of Player 2, Player 1 selects the Top Strategy with probability P_t, and the Bottom Strategy with probability $P_b = 1 - P_t$. As such, Player 1 is essentially choosing between the two gambles T and B:

$$T = \begin{cases} a & \text{w.p. } P_l \\ b & \text{w.p. } 1 - P_l \end{cases} \tag{8}$$

$$B = \begin{cases} c & \text{w.p. } P_l \\ d & \text{w.p. } 1 - P_l \end{cases} \tag{9}$$

with gambles T, B corresponding to choosing the Top Strategy and the Bottom Strategy, respectively, and Player 2 is essentially choosing between the two gambles L and R:

$$L = \begin{cases} x & \text{w.p. } P_t \\ y & \text{w.p. } 1 - P_t \end{cases} \tag{10}$$

$$R = \begin{cases} v & \text{w.p. } P_t \\ w & \text{w.p. } 1 - P_t \end{cases} \tag{11}$$

with gambles L, R corresponding to choosing the Left Strategy and the Right Strategy, respectively.

The line of reasoning presented above shows that the problem of strategy selection for a player in 2ONGs can be formally reduced to the problem of deciding between two risky gambles (T, B for Row Player and L, R for Column Player). By the same logic, more generally, the problem of strategy selection for a player in an n-player ONG (with each player having n pure strategies to choose from) can be formally reduced the problem of deciding between n risky gambles.

As evidenced by Eqs. (8-9) depending on the parameter P_l, Player 1's choice between T and B explicitly depends on Player 1's conception of the probability with which Player 2 would select the Left Strategy (i.e., P_l). Likewise, as evidenced by Eqs. (10-11), Player 2's choice between L and R explicitly depends on Player 2's conception of the probability with which Player 1 would select the Top Strategy (i.e., P_t).

As a case-study, in the next section we turn our attention to Prisoner's Dilemma, and we show that, together with the general way of reducing ONGs to risky decision-making discussed above, SbEU can accurately explain cooperation in 2ONPDs, thereby providing a process-level, rational basis for cooperation in 2ONPDs.

4 Cooperation in One-shot, Non-cooperative Prisoner's Dilemma

A wealth of real-life scenarios are modeled as an instance of Prisoner's Dilemma, e.g., conflict of two prisoners independently questioned by the police (Kaminski, 2003), cartel problems (Osborne, 1976), the conflict of two superpowers who engage in a nuclear arms race (Wiesner & York, 1964), doping in sports (Savulescu, Foddy, & Clayton, 2004; Haugen, 2004), and global warming (Milinski et al., 2008).

Although, normatively, one should never cooperate in one-shot, non-cooperative Prisoner's Dilemma games, substantial experimental evidence shows that people typically cooperate in 2ONPDs (e.g., Dawes & Thaler, 1988; Fehr & Gächter, 2000; Keser & van Winden, 2000; Brandts & Schram, 2001).

In a 2ONPD, each player has two strategies to choose from: either to cooperate or to defect. The payoff matrix of a generic 2ONPD is shown in Fig. 2.

		Player 2 (Column Player)	
		Cooperate	Defect
Player 1 (Row Player)	Cooperate	r, r	v, t
	Defect	t, v	p, p

Figure 2: Payoff matrix of a generic 2-player, one-shot, non-cooperative Prisoner's Dilemma (2ONPD), where $t > r > p > v$. For 2ONPDs, the constraint $r > p$ ensures that mutual cooperation is superior to mutual defection, while the constraints $t > r$ and $p > v$ grant that defection is the dominant strategy for both players. Players can either cooperate or defect.

According to the general framework presented in the previous section, assuming that (from the perspective of a player) the other player would cooperate with probability P_c, a player is essentially choosing from the following two risky choices:

$$\text{Cooperate} = \begin{cases} r & \text{w.p. } P_c \\ v & \text{w.p. } 1 - P_c \end{cases}$$

$$\text{Defect} = \begin{cases} t & \text{w.p. } P_c \\ p & \text{w.p. } 1 - P_c \end{cases}$$

According to the normative principle of least-informative priors (i.e., those prior distributions attaining highest entropy), having no priori knowledge of, or any opportunity to learn through interactions about, her opponent—due to the one-shot, non-cooperative nature of the game—it is rationally justified for a player to assume that $P_c = 0.5$. Accordingly, throughout this paper we make the assumption that $P_c = 0.5$.

Recent work has provided mounting evidence suggesting that people often use very few samples in probabilistic judgments and reasoning (e.g., Vul et al., 2014; Battaglia et al. 2013; Lake et al., 2017; Gershman, Horvitz, & Tenenbaum, 2015; Hertwig & Pleskac, 2010; Griffiths et al., 2012; Gershman, Vul, & Tenenbaum, 2012; Bonawitz et al., 2014; Nobandegani et al., 2018; Lieder, Griffiths, Huys, & Goodman, 2018). Consistent with this finding, throughout this paper we assume that a player draws very few samples ($s = 1$; see Eqs. (2-3)) when deciding between cooperation and defection in 2ONPDs—except for Sec. 4.3 in which we directly investigate the effect of number of samples s on cooperation.

Under these justified assumptions (i.e, $s = 1$ and $P_c = 0.5$), in the following two subsections we show that Nobandegani et al.'s (2018) metacognitively-rational model, SbEU, accurately explains how cooperation rate varies depending on the parameterization of a 2ONPD.

4.1 Manipulation of Cooperation Index

Introduced by Rapoport and Chammah (1965), *cooperation index* (CI) is a concrete measure of cooperativeness in 2ONPDs; CI is a property of the experimental task. For a 2ONPD with a generic payoff matrix shown in Fig. 2, CI is given by (Rapoport & Chammah, 1965):

$$\mathrm{CI} = \frac{r-p}{t-v}. \tag{12}$$

As for any 2ONPD holds $t > r > p > v$ (see Fig. 2), it follows that $0 < \mathrm{CI} < 1$. (The latter result follows from having $t - v > r - p > 0$.)

Rapoport and Chammah (1965) experimentally demonstrated a linear relationship between CI and cooperation rate, with people tending to cooperate more as CI increases. Several studies have replicated this finding (e.g., Steele & Tedeschi, 1967; Vlaev & Chater, 2006).

Next, we show that SbEU can remarkably accurately account for this finding. To test how the cooperation rate predicted by SbEU changes as CI increases, we use nine representative 2ONPD games from Vlaev and Chater (2006, Table 1) which allow us to systematically vary CI equidistantly between 0.1 and 0.9. Recall that $0 < \mathrm{CI} < 1$. We simulate $N = 100,000$ participants, with $s = 1$ and $P_c = 0.5$.

As Fig. 3 demonstrates, there is a significant positive, linear relationship between CI and the cooperation rate predicted by SbEU (Pearson's $r = .9998$, Kendall's $\tau = 1$, Spearman's $\rho = 1$, Ps $< 10^{-5}$).

In the next subsection, we directly compare the cooperation rate predicted by SbEU and human data.

4.2 Manipulation of Defection Payoff p

In a recent experiment investigating the effect of manipulation of defection payoff (i.e., parameter p; see Fig. 2) on cooperation, Engel and Zhurakhovska (2016) presented participants with eleven 2ONPDs; across these stimuli, they systematically varied parameter p while keeping the other parameters fixed (for experimental stimuli see Engel and Zhurakhovska, 2016, Sec. 3).

Fig. 4 shows that SbEU can remarkably accurately account for Engel and Zhurakhovska's (2016) observed cooperation rates, explaining 98% of the variance in the experimental data (Pearson's $r = .9906$, Kendall's $\tau = .9909$, Spearman's $\rho = .9977$, Ps $< .001$). In Fig. 4, we simulate $N = 100,000$ participants, with $s = 1$ and $P_c = 0.5$.

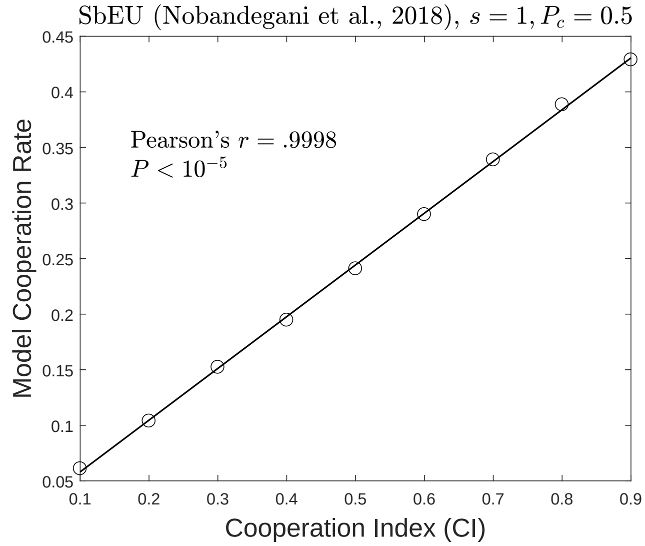

Figure 3: SbEU (Nobandegani et al., 2018) can accurately simulate the linear relationship between CI and cooperation rate, experimentally demonstrated by Rapoport and Chammah (1965).

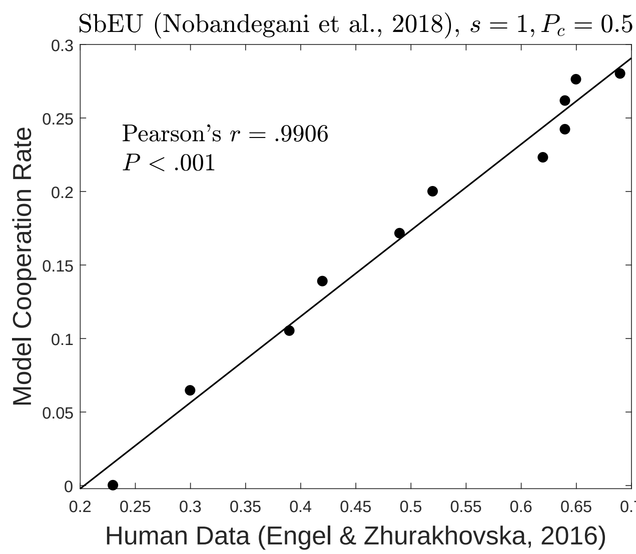

Figure 4: SbEU (Nobandegani et al., 2018) simulates Engel and Zhurakhovska's (2016) experimental data on the effect of manipulation of defection payoff (i.e., parameter p) on cooperation in 2ONPDs.

4.3 The Predictive Relationship between Number of Samples s and Cooperation

In a recent study, Kanazawa and Fontaine (2013) experimentally investigated the effect of general intelligence (measured by a Raven's-type nonverbal test of general intelligence) on cooperation in 2ONPDs, showing that individuals with higher general intelligence are less likely to cooperate.

In this section we investigate the predictive relationship between the number of samples s and cooperation rate in 2ONPDs. In the context of SbEU, we operationalize the well-supported assumption that people with higher general intelligence typically enjoy more cognitive resources, e.g. working memory (e.g., Colom, Jung, & Haier, 2007; Colom et al., 2008, Burgess, Gray, Conway, & Braver, 2011) by positing that these individuals tend to draw more samples.

Consistent with Kanazawa and Fontaine's (2013) finding, SbEU predicts that cooperation rate should decrease as the number of samples s increases; see Fig. 5. In Fig. 5, we adopt the Kanazawa and Fontaine's (2013) specific PD problem given to the subjects (a 2ONPD with $r = 3, v = 0, t = 5, p = 1$), and simulate $N = 100,000$ participants with $P_c = 0.5$.

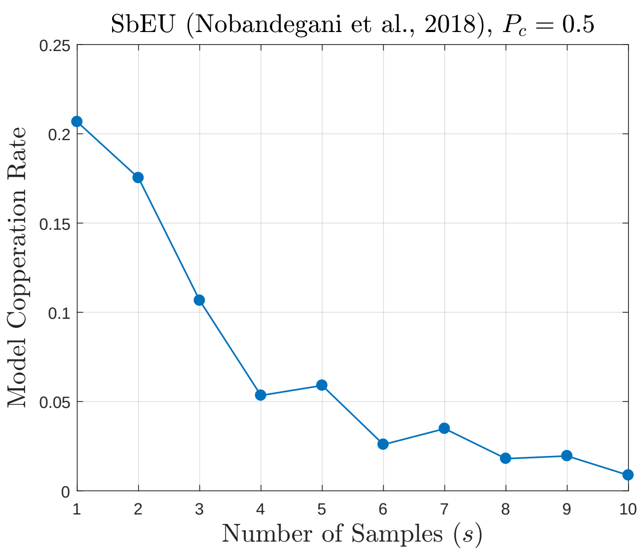

Figure 5: SbEU (Nobandegani et al., 2018) predicts that cooperation rate should decrease as the number of samples s increases, consistent with the experimental findings of Kanazawa and Fontaine (2013).

Importantly, SbEU's prediction depicted in Fig. 5 is supported by substantial evidence revealing that, in the context of 2ONPDs, deliberation (which can be readily operationalized in terms of drawing more samples) leads to increased defection rate, thus bringing behavior closer to the prescriptions of the normative standards of game theory (e.g., Rand, 2016).

4.4 Manipulation of P_c

Shafir and Tversky (1992) examined cooperation rates in a well-known variant of 2ONPD: In some trials, participants were told what the other player was doing. Unsurprisingly, when participants were told that the other person decided to defect, then their probability to cooperate was 0.03; and when they were told that the other person decided to cooperate, then their probability to cooperate was 0.16. However, in trials (within participants design) when participants were not told what the other person did, the probability to cooperate raised to 0.37. This pattern of responding has been independently replicated several times (e.g., Busemeyer, Matthew, & Wang, 2006; Croson, 1999; Li & Taplin, 2002), and has thus far remained a puzzle for optimal decision theorists to explain.

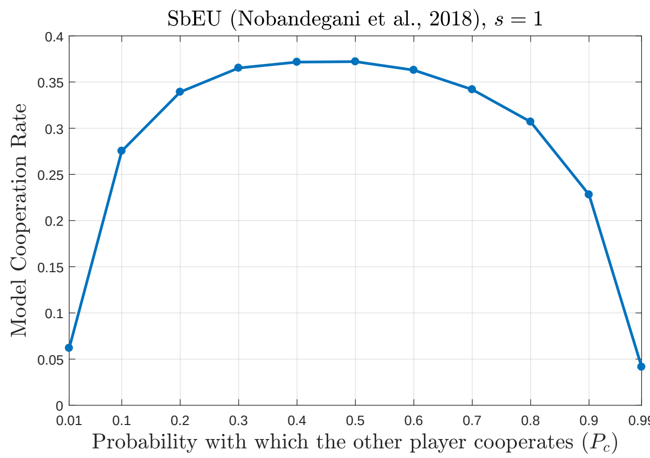

Figure 6: SbEU (Nobandegani et al., 2018) provides a resource-rational, process-level explanation of the puzzling finding of Shafir and Tversky (1992). This finding thus far has defied any (Bayesian) rational explanation. We simulate $N = 100,000$ participants, with $s = 1$. We use a representative 2ONPD game from Shafir and Tversky (1992, Fig. 28.2) with the following parameters: $r = 75, v = 25, t = 85, p = 30$.

The present study offers one, and thus far the only, (Bayesian) rational process-level explanation of this puzzle. As Fig. 6 shows, SbEU predicts that a participant should have only a minuscule tendency to cooperate when the other player is known to either fully cooperate or defect. However, consistent with Shafir and Tversky's (1992) finding, SbEU predicts that participants should have a substantially greater tendency to cooperate when they are maximally uncertain about what the other player would do. As such, SbEU provides a rational explanation of a clear violation of the *law of total probability* in 2ONPDs, as demonstrated by Shafir and Tversky (1992). According to the law of total probability (Durrett, 2010), the cooperation rate under the condition that the opponent's choice is unknown must fall between the cooperation rates observed under the two extreme conditions: full cooperation and full defection.

5 General Discussion

Despite its solid normative ground, NE has failed to provide a satisfying descriptive account of human behavior in many game-theoretic settings. In this work, we focus on a welldocumented, yet puzzling, deviation from NE: cooperation in 2-player, one-shot, non-cooperative, Prisoner's Dilemma games (2ONPDs). By way of introducing a general framework allowing us to conceptualize strategy selection in one-shot, non-cooperative games (ONGs) as the classical problem of decision-making under risk, we investigate whether (seemingly irrational) cooperation in 2ONPDs could be understood as an optimal behavior with the mind acting as a miser.

To our knowledge, our work provides the first (but, admittedly, preliminary) demonstration of how cooperation can arise from *purely selfish*, expected-utility maximization under cognitive limitations. Our findings challenge the widespread view that observed cooperation in 2ONPD games is primarily due to "cooperation bias" in humans, and are supported by recent experimental findings revealing little evidence for such cooperation bias (Pothos et al., 2011). As such, our work refutes the (very intuitive) widely-accepted conclusion that "If players are egoists, cooperation will not be observed in one-shot PD games" (Cooper et al., 1996).

Concretely, in this work we show that the Nobandegani et al.'s (2018) metacognitively-rational process model, SbEU, provides a resource-rational mechanistic explanation of cooperation in 2ONPDs, and offers an accurate quantitative account of how cooperation rate varies depending on the parameterization of a 2ONPD. Furthermore, by operationalizing higher intelligence in terms of drawing a larger number of samples in the available time, our work predicts that more intelligent individuals should tend to cooperate less, fully consistent with recent experimental findings (Kanazawa & Fontaine, 2013).

Shafir and Tversky's (1992) paradoxical finding on the violation of the law of total probability in 2ONPDs has resisted explanation by a model governed by classical probability theory (CPT) for nearly three decades. Interestingly, this paradoxical finding has been recently taken as strong evidence for quantum-probability models of cognition (e.g., Pothos & Busemeyer, 2009; Pothos & Busemeyer, 2013). Our work offers the first, and thus far the only, CPT-based explanation of Shafir and Tversky's (1992) paradoxical finding. As such, our work corroborates the view that decision-making behaviors that appear to be inconsistent with CPT, might after all be reconcilable with CPT when analyzed from an algorithmic perspective acknowledging cognitive limitations.

Being primarily inspired by the experimental finding that deliberation leads to a marked increase in defection rate in 2ONPDs (e.g., Rand, 2016), and applying a dual-process lens to cooperation in 2ONPDs, some researchers have recently argued that intuition favors cooperation while deliberation promotes selfishness (e.g., Rand, Greene, & Nowak, 2012; Rubinstein, 2007; Rand, 2016). Our work offers a completely new way of understanding this experimental finding—both qualitatively and quantitatively.

On the quantitative front, in sharp contrast to a dual-process perspective, our work presents the first, and thus far the only, *single-process* model of cooperation in 2ONPDs, providing a resource-rational mechanistic explanation of why deliberation leads to increased defection. According to our work, it is the optimal use of limited cognitive resources that underlies deliberation promoting selfishness in 2ONPDs. Relatedly, our recent work on modeling fairness in the Ultimatum Game (UG) also supports this view (Nobandegani, Destais, & Shultz, in prep).

On the qualitative side, our work offers a radically different interpretation of cooperation in 2ONPDs than the one provided by the classical dual-process account. From a dual-process perspective, intuition (moderated by System 1) is good and cooperative while deliberation (moderated by System 2) is evil and uncooperative. However, according to our singe-process model (SbEU; Nobandegani et al., 2018), a boundedly-rational agent that selfishly maximizes its expected utility while optimally using its limited cognitive resources should show the highest cooperation rate as an intuitive response, with cooperation rate declining with deliberation. As such, according to our work, humans' intuitive response being to cooperate in 2ONPDs, is still, quite counter-intuitively, the effect of selfishly maximizing expected utility while optimally using limited cognitive resources.

Our work contributes to an emerging line of work attempting to explain human cognition as an optimal use of limited cognitive resources (*rational minimalist program*, Nobandegani, 2017; Griffiths, Lieder, & Goodman, 2015), thereby demonstrating that a wide range of human behaviors are rational, provided that the computational and cognitive limitations of the mind are taken into consideration (Simon, 1957).

By demonstrating that SbEU, a recently proposed metacognitively-rational process model of risky choice (Nobandegani et al., 2018), can quantitatively account for ostensibly irrational cooperation rates in 2ONPDs, our work bridges between two related, but distinct, areas of research: game-theoretic decision-making and risky decision-making. As such, the work presented here brings us a step closer to developing a unified, mechanistic account of human decision-making.

Recent work has shown that SbEU can account for the St. Petersburg paradox, a centuries-old paradox in human decision-making (Nobandegani, da Silva Castanheira, Shultz, & Otto, 2019b), and has experimentally confirmed a counterintuitive prediction of SbEU: Deliberation leads people to move from one well-known bias, framing effect, to another well-known bias, the fourfold pattern of risk preferences (da Silva Castanheira; Nobandegani, & Otto, 2019). An important line of future work would be to investigate whether SbEU could also serve as a resource-rational process-level account of contextual effects in multi-attribute decision-making (e.g., the attraction, similarity, and compromise effects), thus bringing us another step closer to developing this unified account of decision-making.

Acknowledgments This work is supported by an operating grant to TRS from the Natural Sciences and Engineering Research Council of Canada. We would like to thank Constance

Destais, Ashley Stendel, Marcel Montrey, and Peter Helfer for helpful comments on an earlier draft of this work.

References

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, *110*(45), 18327–18332.

Bernheim, B. D. (1984). Rationalizable strategic behavior. *Econometrica: Journal of the Econometric Society*, 1007–1028.

Brandts, J., & Schram, A. (2001). Cooperation and noise in public goods experiments: applying the contribution function approach. *Journal of Public Economics*, *79*(2), 399–427.

Burgess, G. C., Gray, J. R., Conway, A. R., & Braver, T. S. (2011). Neural mechanisms of interference control underlie the relationship between fluid intelligence and working memory span. *Journal of Experimental Psychology: General*, *140*(4), 674.

Busemeyer, J. R., Matthew, M. R., & Wang, Z. (2006). A quantum information processing explanation of disjunction effects. In: *Proceedings of the* 28^{th} *Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Colom, R., Abad, F. J., Quiroga, M. Á., Shih, P. C., & Flores-Mendoza, C. (2008). Working memory and intelligence are highly related constructs, but why? *Intelligence*, *36*(6), 584–606.

Colom, R., Jung, R. E., & Haier, R. J. (2007). General intelligence and memory span: evidence for a common neuroanatomic framework. *Cognitive Neuropsychology*, *24*(8), 867–878.

Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1996). Cooperation without reputation: experimental evidence from prisoner's dilemma games. *Games and Economic Behavior*, *12*(2), 187–218.

Croson, R. T. (1999). The disjunction effect and reason-based choice in games. *Organizational Behavior and Human Decision Processes*, *80*(2), 118–133.

da Silva Castanheira, K., Nobandegani, A. S., & Otto, A. R. (2019). Sample-based variant of expected utility explains effects of time pressure and individual differences in processing speed on risk preferences. In *Proceedings of the* 41^{st} *Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Daskalakis, C., Goldberg, P. W., & Papadimitriou, C. H. (2009). The complexity of computing a Nash equilibrium. *SIAM Journal on Computing*, *39*(1), 195–259.

Dawes, R. M., & Thaler, R. H. (1988). Anomalies: cooperation. *Journal of Economic Perspectives*, *2*(3), 187–197.

Durrett, R. (2010). *Probability: Theory and Examples*. Cambridge university press.

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

Fehr, E., & Gachter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, *90*(4), 980–994.

Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, *349*(6245), 273–278.

Gershman, S. J., Vul, E., & Tenenbaum, J. B. (2012). Multistability and perceptual inference. *Neural Computation*, *24*(1), 1–24.

Geweke, J. (1989). Bayesian inference in econometric models using Monte Carlo integration. *Econometrica: Journal of the Econometric Society*, 1317–1339.

Goeree, J. K., & Holt, C. A. (2001). Ten little treasures of game theory and ten intuitive contradictions. *American Economic Review*, *91*(5), 1402–1422.

Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, *7*(2), 217–229.

Griffiths, T. L., Vul, E., & Sanborn, A. N. (2012). Bridging levels of analysis for probabilistic models of cognition. *Current Directions in Psychological Science*, *21*(4), 263–268.

Hammersley, J., & Handscomb, D. (1964). *Monte carlo methods.* London: Methuen & Co Ltd.

Haugen, K. K. (2004). The performance-enhancing drug game. *Journal of Sports Economics*, *5*(1), 67–86.

Hertwig, R., & Pleskac, T. J. (2010). Decisions from experience: Why small samples? *Cognition*, *115*(2), 225–237.

Holt, C. A., & Roth, A. E. (2004). The Nash equilibrium: A perspective. *Proceedings of the National Academy of Sciences*, *101*(12), 3999–4002.

Kahneman, . T. A., D. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

Kaminski, M. M. (2018). *Games prisoners play: The tragicomic worlds of Polish prison*. Princeton University Press.

Kanazawa, S., & Fontaine, L. (2013). Intelligent people defect more in a one-shot prisoners dilemma game. *Journal of Neuroscience, Psychology, and Economics*, *6*(3), 201.

Keser, C., & Van Winden, F. (2000). Conditional cooperation and voluntary contributions to public goods. *Scandinavian Journal of Economics*, *102*(1), 23–39.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, *40*.

Li, S., & Taplin, J. (2002). Examining whether there is a disjunction effect in prisoner's dilemma games. *Chinese Journal of Psychology*, *44*, 25–46.

Lieder, F., Griffiths, T. L., Huys, Q. J., & Goodman, N. D. (2018). Empirical evidence for resource-rational anchoring and adjustment. *Psychonomic Bulletin & Review*, *25*(2), 775–784.

Mailath, G. J. (1998). Do people play Nash equilibrium? lessons from evolutionary game theory. *Journal of Economic Literature*, *36*(3), 1347–1374.

Markowitz, H. (1952). The utility of wealth. *Journal of political Economy*, *60*(2), 151–158.

Milinski, M., Sommerfeld, R. D., Krambeck, H.-J., Reed, F. A., & Marotzke, J. (2008). The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences*, *105*(7), 2291–2294.

Nash, J. F., et al. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, *36*(1), 48–49.

Nobandegani, A. S. (2017). The Minimalist Mind: On Mininality in Learning, Reasoning, Action, & Imagination. McGill University, PhD Dissertation.

Nobandegani, A. S., da Silva Castanheira, K., O'Donnell, T. J., & Shultz, T. R. (2019a). On robustness: An undervalued dimension of human rationality. In *Proceedings of the* 17^{th} *International Conference on Cognitive Modeling.* Montreal, QC.

Nobandegani, A. S., da Silva Castanheira, K., Otto, A. R., & Shultz, T. R. (2018). Over-representation of extreme events in decision-making: A rational metacognitive account. In *Proceedings of the* 40^{th} *Annual Conference of the Cognitive Science Society* (pp. 2391-2396). Austin, TX: Cognitive Science Society.

Nobandegani, A. S., da Silva Castanheira, K., Otto, A. R., & Shultz, T. R. (2019b). A resource-rational process-level account the st. petersburg paradox. In *Proceedings of the* 41^{st} *Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Nobandegani, A. S., Destais, C., & Shultz, T. R. (in prep.). An expectation-based model of fairness in the Ultimatum Game.

Osborne, D. K. (1976). Cartel problems. *The American Economic Review*, *66*(5), 835–844.

Pearce, D. G. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica: Journal of the Econometric Society*, 1029–1050.

Poor, H. V. (2013). *An Introduction to Signal Detection and Estimation*. Springer Science & Business Media.

Pothos, E. M., & Busemeyer, J. R. (2009). A quantum probability explanation for violations of rationaldecision theory. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1665), 2171–2178.

Pothos, E. M., & Busemeyer, J. R. (2013). Can quantum probability provide a new direction for cognitive modeling? *Behavioral and Brain Sciences*, *36*(3), 255–274.

Pothos, E. M., Perry, G., Corr, P. J., Matthew, M. R., & Busemeyer, J. R. (2011). Understanding cooperation in the prisoners dilemma game. *Personality and Individual Differences*, *51*(3), 210–215.

Rand, D. G. (2016). Cooperation, fast and slow: Meta-analytic evidence for a theory of social heuristics and self-interested deliberation. *Psychological Science*, *27*(9), 1192–1206.

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

Rapoport, A., Chammah, A. M., & Orwant, C. J. (1965). *Prisoner's dilemma: A study in conflict and cooperation* (Vol. 165). University of Michigan press.

Rubinstein, A. (2007). Instinctive and cognitive reasoning: A study of response times. *The Economic Journal*, *117*(523), 1243–1259.

Savulescu, J., Foddy, B., & Clayton, M. (2004). Why we should allow performance enhancing drugs in sport. *British Journal of Sports Medicine*, *38*(6), 666–670.

Scholten, M., & Read, D. (2014). Prospect theory and the forgotten fourfold pattern of risk preferences. *Journal of Risk and Uncertainty*, *48*(1), 67–83.

Shafir, E., & Tversky, A. (1992). Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive Psychology*, *24*(4), 449–474.

Simon, H. A. (1957). *Models of Man*. Wiley.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207–232.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323.

Vul, E., Goodman, N., Griffiths, T. L., & Tenenbaum, J. B. (2014). One and done? optimal decisions from very few samples. *Cognitive Science*, *38*(4), 599–637.

Wiesner, J. B., & York, H. F. (1964). National security and the nuclear-test ban. *Scientific American*, *211*(4), 27–35.