Universitat Pompeu Fabra Barcelona
Department of Economics and Business

Economics Working Paper Series

Working Paper No. 1849

Coordination and sophistication

Larbi Alaoui, Katharina A. Janezic,
and Antonio Penta

November 2022

Coordination and Sophistication*

Larbi Alaoui† Katharina A. Janezic‡ Antonio Penta§

April 2022

Abstract

How coordination can be achieved in isolated, one-shot interactions without communication and in the absence of focal points is a long-standing question in game theory. We show that a cost-benefit approach to reasoning in strategic settings delivers sharp theoretical predictions that address this central question. In particular, our model predicts that, for a large class of individual reasoning processes, coordination in some canonical games is more likely to arise when players perceive heterogeneity in their cognitive abilities, rather than homogeneity. In addition, and perhaps contrary to common perception, it is not necessarily the case that being of higher cognitive sophistication is beneficial to the agent: in some coordination games, the opposite is true. We show that subjects' behavior in a laboratory experiment is consistent with the predictions of this model, and reject alternative coordination mechanisms. Overall, the empirical results strongly support our model.

Keywords: coordination – cognitive cost – sophistication – strategic reasoning – value of reasoning

JEL Codes: C72; C91; C92; D80.

1 Introduction

Individuals are often faced with situations in which they must attempt to coordinate despite having very little information about their opponents, or on past behavior. In such settings, whether coordination can be achieved at all has been an important open question in game theory. One proposed mechanism uses a notion of focal points (Schelling (1960)), which depends on the existence of a shared culture, since there must be a common view

*We thank several seminar and conference audiences where these results were presented. Special thanks go to Johannes Abeler, Vince Crawford, David Gill, Lukas Hoesch, Nagore Iriberri, Willemien Kets, Eugenio Proto, Victoria Prowse, Ariel Rubinstein, Aldo Rustichini, Jakub Steiner and Séverine Toussaert. The BSE acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Programme for Centres of Excellence in R&D (CEX2019-000915-S), Larbi Alaoui from the Spanish Ministry of Science and Innovation under projects ECO2014-56154-P, PGC2018-098949- B-I00 and grant RYC-2016-21127, and Antonio Penta from the ERC Starting Grant #759424.

†Universitat Pompeu Fabra, BSE and UM6P (AIRESS). E-mail: larbi.alaoui@upf.edu

‡University of Oxford. E-mail: katharina.janezic@economics.ox.ac.uk

§ICREA, Universitat Pompeu Fabra, BSE and TSE. E-mail: antonio.penta@upf.edu

	Bach	*Stravinsky*
Bach	$x, 50$	$0, 0$
Stravinsky	$0, 0$	$50, x$

Figure 1: The 'canonical' BoS, with $x > 50$

concerning which points are focal. In practice, however, and especially when agents face novel strategic situations, these conditions are often not met. Then, players can only resort to their introspective reasoning, and it is again unclear how, or even whether, coordination can be achieved on a purely *eductive* basis (cf. Binmore (1987, 1988)).[1] This is the case especially when the coordination problem is accompanied by an element of conflict, as exemplified by the canonical Battle of the Sexes (BoS, see Figure 1): in such situations, overcoming the coordination problem also requires a solution to the bargaining problem that is implicit in the equilibrium selection.

In this paper, we introduce a novel mechanism for how coordination can be achieved even in the absence of focality. This mechanism is based on the fact that, even in such settings, individuals may form beliefs over each other's cognitive sophistication (cf. Proto, Rustichini, and Sofianos (2019, 2021) and Lambrecht, Proto, Rustichini, and Sofianos (2022)). We will show that, under a broad class of reasoning processes and in the absence of focal points, coordination can be achieved if players' reasoning responds to incentives (cf. Alaoui and Penta (2016, 2022)) and if they view each other as having *different* cognitive abilities. Perhaps surprisingly, our model predicts that it is *heterogeneity* between players, rather than homogeneity, that favors coordination in settings with inherent strategic uncertainty. This mechanism therefore brings to light a very different aspect of coordination, compared to that based on a shared culture and focality, where the logic of coordination is associated with some form of homogeneity among players (cf. Kets and Sandroni (2019, 2021), Kets, Kager, and Sandroni (2022) and Kets (2022)).

To fix ideas, consider the canonical BoS in Figure 1, in which two players both prefer coordinating on the same location (say, between a Bach or a Stravinsky concert), but have different preferences over which place to coordinate on. If neither location is focal, and if there is no communication nor experience of past play, then it is difficult to see how these agents can coordinate. Now suppose that the players have beliefs over each other's 'cognitive sophistication', in the sense of how costly it is for them to reason about what the other might do. Consider the following two situations. In case (i), the agents view each other to be of similar sophistication; in case (ii), they (commonly) believe that one player is of higher sophistication than the other. In our model, these beliefs have clear implications on the likelihood of coordination, and on where the agents will coordinate.

In particular, in the situations described above, our model predicts higher coordination

[1] The term 'eductive' is introduced by Binmore (1987, 1988), to refer to the rationalistic, reasoning-based approach to the foundations of solution concepts. The 'eductive approach' is contrasted with the 'evolutive approach', in which solution concepts are interpreted as the steady states of an underlying learning or evolutionary process.

in case (ii), when players believe that they have different sophistication, than in case (i), when they believe they are of similar sophistication. Furthermore, the model predicts that players in the *canonical* BoS are more likely to coordinate on the equilibrium that is most favorable to the player who is believed to be *less* sophisticated. That is, when players face strategic uncertainty in the *canonical* BoS, we predict that being perceived to be *more* sophisticated is a disadvantage (as we discuss below, however, the opposite is true in other games). Finally, and in contrast with what one might expect, payoff transformations that exacerbate the disagreement between players in the BoS (while preserving the symmetry of the game – e.g., increasing x in Figure 1), do not reduce coordination, and may in fact favor its occurrence.

An important feature of our results is that they follow from minimal assumptions on players' form of reasoning, on their costs, and on their beliefs about each other. The key assumption concerning beliefs is that players agree (rightly or not) on who has relatively lower costs of reasoning. Importantly, the results hold even if beliefs about their opponents' costs, or their very form of reasoning, are incorrect. As for the form of reasoning, our model is essentially unrestricted, and it accommodates an equilibrium selection procedure, as perhaps someone trained in game theory would follow, or a level-k form of reasoning (e.g., Nagel (1995), Crawford et al. (2013)), or entirely different ways of reasoning altogether. The only requirement is that it has to be *responsive*, in that thinking more never loses the potential to change the player's understanding of the situation. This is a natural way of formalizing the inherent strategic uncertainty that arises in the absence of focal points, and formally captures the sense in which our model is complementary to that of Kets and Sandroni (2019, 2021). Another important aspect of our model is that it builds on an existing approach that has both theoretical and extensive empirical support (see the general axiomatic framework of Alaoui and Penta (2022) and the experimental evidence on the endogenous level-k model in Alaoui and Penta (2016) and Alaoui, Janezic, and Penta (2020)).[2] Under this overall approach, each player tries to understand the other player's reasoning, and where his reasoning may have stopped. Based on his own reasoning and his belief about the opponent's reasoning, the individual chooses his action. His behavior is thus jointly determined by the stakes of the game (which determine the value), his own sophistication (cost of reasoning), and his beliefs about the opponent's sophistication.

Returning to the BoS game above, the logic underlying our main results is the following. Each player reasons through the game according to his form of reasoning (as discussed above, we allow for very general paths of reasoning). We show that, if x is sufficiently high in the parametrization of Figure 1, then a player's reasoning is more likely to eventually end with the action associated with his favorite equilibrium (this is *not* true for other

[2] This approach has also been shown to be consistent with the experimental results in Goeree and Holt (2001) and Esteban-Casanelles and Gonçalves (2020), with the experiments on response time and attention allocation by Alós-Ferrer and Buckenmaier (2021), and others. For further discussions, see Alaoui and Penta (2022) and Kagel and Penta (2021). See also Gill and Prowse (2022) on strategic complexity and the value of thinking in a setting with response times, and for the importance of strategic sophistication and lifetime outcomes see Fe, Gill, and Prowse (2022).

coordination games, as we explain below, or for reasoning processes that 'stabilize' (cf. Kets and Sandroni (2019, 2021)), or for those that arise in the presence of focal points). Now consider again the two cases spelled out above, in which the players commonly believe that they are (i) of similar cognitive sophistication or (ii) of different sophistication. In case (i), both players play according to their *own* understanding for the situation, and will only coordinate if, by coincidence, their own reasoning processes are aligned in a fortunate way. In fact, if x is high enough, then we predict that they would actually *mis*coordinate, each attempting to coordinate on their favorite equilibrium. In case (ii), the player perceived to be less sophisticated still plays according to his own understanding of the situation, but the other thinks he has gone deeper in his reasoning than his opponent. He therefore plays according to where he believes the opponent has stopped reasoning, rather than according to where his own reasoning has stopped. If each player's favorite equilibrium is sufficiently preferred (i.e., if x is high enough), then more sophisticated player will expect his opponent to choose his favorite action, and hence best responds to that. As a consequence, the players are more likely to coordinate, and they will do so on the preferred equilibrium of the player perceived to be less sophisticated. Note that at no point does this logic rely on the players' perceptions about one another to be correct, provided that they commonly agree on their relative sophistication.

After introducing our model and formally providing our theoretical results, we conduct a laboratory experiment to test the predictions. First, subjects take a test of strategic sophistication, and are labeled according to their scores. The higher and lowest scoring subjects play both against their own and against the other label in the *BoS* game as in Figure 1, with the x parameter set equal to 51 and 70, which we refer to respectively as 'low payoff' and 'high payoff' treatments. Subjects also play other games, that we discuss below. The empirical findings are in line with the predictions of the model. Specifically, in the *canonical BoS*, we find that: (i) high label subjects concede more against low than against high; (ii) this effect is more pronounced for the high payoff treatments than low payoff treatments; (iii) low label subjects play in a similar manner against low as against high, for both payoffs; (iv) there is more coordination when playing against the other label than against their own; (v) the increased coordination occurs on the low label's favorite equilibrium. Put together, this set of results lends support to the empirical relevance of the mechanism introduced by our model.

A question that can arise is whether other mechanisms might bring about similar behavioral patterns. For instance, it might seem that the argument provided above to explain the model's predictions is suggestive of a form of *first-mover advantage* for the low type, in the sense that it is as if the low type 'commits' to stop reasoning first, and at his preferred action profile, while the high type then concedes. This analogy, however, does not adequately capture the logic of our model. To illustrate the difference, we propose another coordination game, which we refer to as the *modified BoS* (see Figure 2), in which our model delivers the opposite prediction to the one obtained from the 'first mover'

	Bach	*Stravinsky*
Bach	$70, 50$	$-x, -x$
Stravinsky	$0, 0$	$50, 70$

Figure 2: A *modified* BoS Game, with $x \geq 0$.

argument. In this game, the logic of our model is identical to that above, except that the stopping criteria based on the Alaoui and Penta (2022) axiomatization in this case induce players' reasoning *not* to stop at the low type's favorite equilibrium, but rather at the high type's. For this reason, unlike the canonical BoS, in the *modified BoS* we predict that the high type concedes *less* against the low type than against their own type. As for the low type, as above, we expect them to play similarly against both opponent types.

A second natural conjecture is that the *asymmetry* in player labels itself helps achieve more coordination. That is, we could hypothesize that subjects coordinate on the equilibrium that favors one player or the other, depending on whether they consider the high or the low label to be the focal one. With this view, however, it would be difficult to explain how coordination favors the low label subjects in the canonical BoS, and the high label in the modified BoS, as predicted by our theory.

Thus, the predictions of our model are clearly distinct from both the 'label focality' idea and from the interpretation that it produces a first-mover advantage for the low label, and hence these alternative explanations can be tested against the predictions of our model. In our experiment, we do this using a version of the *modified BoS* and we find that the results are neither in line with the view that low types obtain a first-mover advantage nor with the notion of 'label focality'. Lastly, we consider a Stag Hunt game and an Asymmetric Matching Pennies game and again find consistency between the results and our predictions, for both games.

In brief, we show that a cost-benefit approach to reasoning in strategic settings delivers sharp theoretical predictions that address one of the central issues of game theory. The experimental results are in line with our model's predictions, lending further support to the mechanism introduced here.

2 Model

In this section we introduce a model of stepwise reasoning and deliberation, for general two-player games with complete information, $G = (A_i, u_i)_{i=1,2}$, where A_i denotes the set of actions of player $i \in \{1, 2\}$, with typical element a_i, and $u_i : A_1 \times A_2 \rightarrow \mathbb{R}$ denotes players i's payoff function. Our leading example in this section, which will also form the center of our experimental analysis, will be the *canonical* Battle of the Sexes (BoS) game, with payoffs parameterized by $r \in \mathbb{R}$, $r \geq 1$:

	W_2	B_2
B_1	$r, 1$	$0, 0$
W_1	$0, 0$	$1, r$

Figure 3: Battle of the Sexes Game

Player 1 prefers to coordinate on (B_1, W_2) while Player 2 prefers to coordinate on (W_1, B_2) (the labeling of the actions denote, respectively, the 'best' and 'worst' equilibrium action for that player). If none of the actions are salient in some way, then there are no focal points, and game theory does not provide any guidance as to how coordination can be achieved, if at all. Our focus here will be precisely on this case.

We assume that, in their deliberation, both players follow a stepwise reasoning process, and that the number of steps of reasoning they perform is driven by a cost-benefit analysis, which trades off cognitive costs with some notion of *value of reasoning* that is related to the game's payoffs. The model is based on the axiomatic foundation in Alaoui and Penta (2022), and covers a wide variety of stepwise reasoning processes, which include (but are not limited to) the model of *endogenous level-k reasoning* of Alaoui and Penta (2016) and Alaoui, Janezic, and Penta (2020). Given their understanding of the game, players form beliefs about the opponent's understanding, and choose the action that maximizes their expected utility, given their beliefs about the opponent that have resulted from their reasoning. We introduce next the model of such a deliberation process, introducing its main elements in the following order: the path of reasoning; the stopping rule; and the resulting beliefs and choice. Then we explore some predictions of the model in the canonical BoS game, and illustrate the logic of the key *eductive coordination* mechanism that is the object of our experimental investigation.

2.1 The 'Path of Reasoning'

Fix a two-player game with complete information, $G = (A_i, u_i)_{i=1,2}$. For each player i, considered in isolation, his stepwise reasoning process is described by a sequence $\{(a_i^{i,k}, a_j^{i,k})\}_{k\in\mathbb{N}}$, which we refer to as the **path of reasoning**, where for each k, $a_j^{i,k} \in A_j$ represents i's best conjecture, at step k, about the behavior of an opponent that has taken at least that many steps of reasoning. We let $a_i^{i,k} \in BR_i(a_j^{i,k})$, denote his best response to that conjecture, where $BR_i : A_j \rightrightarrows A_i$ denotes player i's pure-action best reply correspondence, defined as $BR_i(a_j) := \arg\max_{a_i\in A_i} u(a_i, a_j)$ for all $a_j \in A_j$.[3]

The path of reasoning $\{(a_i^{i,k}, a_j^{i,k})\}_{k\in\mathbb{N}}$ represents the sequence of conjectures and choices that the agent could potentially consider in his reasoning and deliberation process. As we will formalize below, however, individuals in our model will not reason indefinitely.

[3] The model can be extended to non-degenerate conjectures, of the form $\alpha_j^{i,k} \in \Delta(A_j)$. For simplicity, however, we abstract from this possibility in the introduction of the baseline model, and only focus on degenerate conjectures of the form $a_j^{i,k} \in A_j$. We will discuss the case of non-degenerate conjectures below.

Rather, we view reasoning as costly, and at any given step k players may well decide that it is not worth continuing to reason. If the agent stops reasoning at some step $\hat{k}$, then he may either choose the current action $a_i^{i,\hat{k}}$ or, if he thinks that the opponent stopped reasoning at some lower $k < \hat{k}$, he may choose the corresponding $a_i^{i,k}$ which is optimal given such belief (we will return to the formalization of how the choice is made in the next subsection). We note that the predictions of the model that we will analyze apply to a broad class of paths of reasoning $\{(a_i^{i,k}, a_j^{i,k})\}_{k \in \mathbb{N}}$. Here, however, we discuss a few benchmark examples:

1. Deliberation Over Equilibria: One natural form of reasoning is for a player to progressively understand the equilibria of the game, and deliberate over which one to play. This form of reasoning corresponds to the case in which the path of reasoning also satisfies the condition $a_j^{i,k} \in BR_i(a_i^{i,k})$ for every k. In the BoS game, for instance, players could understand that the possible (pure) equilibria are (B_1, W_2) and (W_1, B_2). At any given step, a player may think that the opponent is trying to coordinate on one equilibrium or the other. It may be that as he thinks more, he remains convinced of this equilibrium, and then his path of reasoning 'stabilizes' at such a profile. Alternatively, it may be that as he thinks more, his reasoning leads him away from that equilibrium to another. In the BoS game, this involves shifting between (B_1, W_2) and (W_1, B_2) as more steps are taken.

2. Level-k Reasoning: Another natural form of reasoning occurs in the form of level-k thinking introduced by Nagel (1995) (see also Crawford, Costa-Gomes, and Iriberri (2013) and references therein). This form of reasoning obtains letting player i's conjecture over the opponent's action at step k be equal to the action of an opponent of level (k-1). Formally, for each $k = 1, 2, ...$, this form of reasoning is such that $a_j^{i,k} = a_j^i\left(k-1\right)$ and $a_l^i\left(k\right) = BR\left(a_{-l}^i\left(k-1\right)\right)$ for each $l \in \{1, 2\}$, where $a^i\left(0\right) = \left(a_1^i\left(0\right), a_2^i\left(0\right)\right)$ is an arbitrary level-0 *anchor*. Note that, for this specific model, if the anchor $a^i\left(0\right)$ is a Nash equilibrium $a^* \in A$, then the path of reasoning is *constant*, in the sense that $a^i\left(k\right) = a^*$ for all k: in this case, in his deliberation player i only contemplates playing the action $a_i^{i,k} = a_i^*$ at every step k. Thus, within the level-k mode of reasoning, the case of a Nash equilibrium anchor can be thought of as a situation in which player i has the initial 'impulse' of playing a^*, and further reasoning about mutual best replies does not challenge such initial disposition. If, in contrast, $a\left(0\right)$ is not an equilibrium, then $\left(a^i(k)\right)$ will not be constant, and may converge or keep cycling. In the BoS, for instance, if $a^i(0) \in \{(B_1, B_2), (W_1, W_2)\}$, then $a_i^i\left(k\right)$ will keep *cycling* between B_i and W_i, which may be taken to represent a situation of a player which, not believing in a focal point, becomes aware of the coordination problem and hence wonders, throughout his path of reasoning and until he stops and makes a choice, over which outcome he should try to coordinate on.

In general, i's path of reasoning could be *'absorbing'*, in the sense that $a_i^{i,k}$ no longer changes past a certain step $k \geq 0$, or it could be *'responsive'*, in the sense that it does not remain stuck at any one best action. For instance, in the case of the 'deliberation over equilibria' mode of reasoning, this would occur if the reasoning does not *ever* stabilize on any one equilibrium. In the case of level-k reasoning, this would be the case if the anchor is a non-Nash equilibrium (i.e., either $\{(B_1, B_2), (W_1, W_2)\}$ in the BoS game). Formally:

Definition 1 *A path of reasoning* $\{(a_i^{i,k}, a_j^{i,k})\}_{k\in\mathbb{N}}$ *of player* i *is* ***absorbing*** *if there exists a* $\bar{k} \geq 0$ *such that, for all* $k > \bar{k}$*,* $a_i^{i,k} = a_i^{i,k+1}$*. A path of reasoning of player* i *is* ***responsive*** *if it is not absorbing.*

If i's path is **absorbing**, then reasoning has no effect past step $\bar{k}$ after which it no longer changes. In the case where such $\bar{k}_i = 0$, reasoning plays no role in changing the player's mind. If both players have *the same* absorbing path of reasoning, with $\bar{k} = 0$ for each, then effectively there is a *focal* action profile, that is shared by the two players, on which they agree. Any possible coordination would thus be due to this focality, and not to their reasoning. Since it is reasoning and not focality that is at the center of our analysis, the bite of our model will be for **responsive** paths.

Other forms of reasoning, however, may be absorbing, but only for some 'high' $\bar{k} > 0$. For instance, Kets and Sandroni (2019, 2021)'s *introspective equilibrium* (see also Kets, Kager, and Sandroni (2022)), describe a reasoning process in which the path of reasoning is generated by a chain of best-responses similar to level-k, but in which players may be of different *types*, each with a possibly different *anchor* (what they call *impulse*). Depending on the type space (which specifies players' types, beliefs, and impulses), and on the payoff of the game, the iteration of the best replies may either converge or not. When such an iteration converges, then it forms an *introspective equilibrium*; otherwise, an introspective equilibrium does not exist for that specific combination of game and type space. From this viewpoint, one can regard the focus of our analysis also as complementary to Kets and Sandroni's: while introspective equilibrium is defined by reasoning processes that converge – and, hence, by paths of reasoning that are *absorbing* – we focus instead on paths of reasoning that remain *responsive*. This property is thus best thought of as one way to capture a situation in which, if players could potentially reason indefinitely (as they do in the Kets and Sandroni (2019, 2021) papers, since k is taken to infinity, there), they would potentially never stop questioning their earlier conclusions. In this sense, responsive paths of reasoning distill the ultimate dilemma in a coordination problem, when no focal points or other fixed point logic can unambiguously pin down a single action profile.

As mentioned above, however, we do not assume that players reason indefinitely. Rather, we view reasoning as costly, and players may well decide, consciously or not, that it is not worth continuing reasoning. In what follows, the main factor is that, all else being the same, a more sophisticated player will stop reasoning at a higher step k than a less sophisticated player. We first explain what leads the agents to stop, based on their cost and value of reasoning for the game in question, and then discuss the agents' beliefs over their opponents. Taken together, the two will determine players' behavior.

2.2 Stopping rule

Player i has **value of reasoning** $v_i(k)$ and a **cost of reasoning** $c_i(k)$ associated with each step of reasoning $k > 0$, where $v_i(k)$ and $c_i(k)$ represent, respectively i's value and cost of doing the k-th round of reasoning, given the previous $k-1$ rounds. Costs represent players' cognitive abilities; the value instead only depend on the game's payoffs, such as the r parameter in the BoS game. When deciding whether or not to reason at that step, the agent compares the two, and continues as long as the value of reasoning exceeds the cost of reasoning, i.e., so long as $v_i(k) \geq c_i(k)$. For future reference, we define a mapping $\mathcal{K} : \mathbb{R}_+^{\mathbb{N}} \times \mathbb{R}_+^{\mathbb{N}} \to \mathbb{N}$ such that, $\forall (c, v) \in \mathbb{R}_+^{\mathbb{N}} \times \mathbb{R}_+^{\mathbb{N}}$,

$$\mathcal{K}(c, v) := \min \{k \in \mathbb{N} : c(k) \leq v(k) \text{ and } c(k+1) > v(k+1)\}, \tag{1}$$

with the understanding that $\mathcal{K}(c, v) = \infty$ if the set in equation (1) is empty. In words, this mapping identifies the first intersection between the value v and the cost c (see Fig. 4). Player i's *cognitive bound* is the value that this function takes at (c_i, v_i):

$$\hat{k}_i = \mathcal{K}(c_i, v_i). \tag{2}$$

To rank players' sophistication, we rank players' costs of reasoning, and refer to cost function c' as 'more sophisticated' than c'' if $c'(k) \leq c''(k)$ for every k (similarly, c' is 'less sophisticated' than c'' if $c'(k) \geq c''(k)$ for all k). Then, for each $c_i \in \mathbb{R}_+^{\mathbb{N}}$, we let $C^+(c_i)$ and $C^-(c_i)$ denote the sets of cost functions that are respectively 'more' and 'less' sophisticated than c_i.

Remark 1 *For any cost of reasoning $c(\cdot)$ and value of reasoning $v(\cdot)$, $\mathcal{K}(v, c) \geq \mathcal{K}(v, c')$ if $c' \in C^-(c)$ and $\mathcal{K}(v, c) \leq \mathcal{K}(v, c'))$ if $c' \in C^+(c)$.*

We assume the following for the cost functions.

Assumption 1 (Cost of Reasoning) *For each i:*

1. *Not thinking is free: $c_i(0) = 0$,*
2. *The cost is increasing: $c_i(k) > c_i(k')$ if $k > k'$.*
3. *Costs are finite: $c_i(k) < \infty$ for all k.*
4. *Costs are not uniformly bounded: $\nexists \bar{c} \in \mathbb{R}$ such that $c_i(k) \leq \bar{c}$ for all k.*

The first property serves as a normalization of the minimal cost of thinking. The content of the second assumption – which could be weakened, as we will discuss – is in essence that of 'theory of mind': for any player, putting himself in the shoes of the opponent putting himself in his own shoes, ..., becomes increasingly difficult. The third assumption ensures that cognitive abilities are not such that they have an absolute limit. This property – which could also be weakened – ensures that the value of reasoning always plays a role. The last assumption rules out the possibility that some high but finite value of reasoning could lead the player to reason endlessly.

We assume that the value of reasoning takes the following form:

$$v_i(k) = \max_{a_j \in A_j} u_i(BR_i(a_j), a_j) - u_i(a_i^{i,k-1}, a_j).$$

This functional form can be interpreted as an extreme form of pessimism over the accuracy of one's current understanding, in that it is as if the agent believes that further reasoning will yield insights which maximize the opportunity cost of stopping. Less extreme forms of the value of reasoning, which for instance consider probability distributions over the opponent's actions which player i may think he would learn about (see, e.g., Alaoui and Penta (2022)), would deliver qualitatively similar implications to the ones we discuss in the following. Hence, while it can be relaxed without significantly affecting the results, the *maximum gain* representation above has the advantage of having no free parameter and thus offering no degrees of freedom. This representation of the value of reasoning will therefore be maintained throughout.

2.3 Beliefs about Others' Reasoning and Choice

The cognitive bound describes the thought process of the agent, but his behavior also depends on his beliefs about his opponent, and particularly over the opponent's cost function. Such beliefs are then used to derive i's beliefs over the opponent's cognitive bound. The *type* of a player is thus described by a pair $t_i = (c_i, c_j^i)$, where c_i represents player i's cost of reasoning, and c_j^i represent his beliefs about player j's cost function.[4] Player i's beliefs about j's cognitive bound will thus be equal to the point where he thinks j has stopped, given his beliefs over his cost of reasoning c_j^i, and taking into account j's value of reasoning, as entailed by i's own understanding of j's reasoning. Formally, let $v_j^i : \mathbb{N} \to \mathbb{R}$ be such that

$$v_j^i(k) = \max_{a_i \in A_i} u_j(BR_j(a_i), a_i) - u_j(a_j^{i,k-1}, a_i).$$

With this notation, we define i's beliefs about j's cognitive bound (given his own bound $\hat{k}_i$, his reasoning path $\{(a_i^{i,k}, a_j^{i,k})\}_{k \in \mathbb{N}}$, and his beliefs about j's cost, c_j^i) as:

$$\hat{k}_j^i = \min \left\{ \hat{k}_i, \mathcal{K}(c_j^i, v_j^i) \right\}. \tag{3}$$

[4]The model can also be extended to include both non-degenerate beliefs about the opponent's cost, as well as higher order beliefs (i.e., i's beliefs about j's beliefs about i's cost, etc.): Following Alaoui and Penta's (2016) EDR model, such belief hierarchies can be modelled through *cognitive type spaces*, which can be used to represent arbitrary belief hierarchies over players' costs (see also Alaoui, Janezic, and Penta (2020)). As we will discuss below, our main results would not be affected by the introduction of non degenerate beliefs, and allowing for more general higher order uncertainty over cost functions.

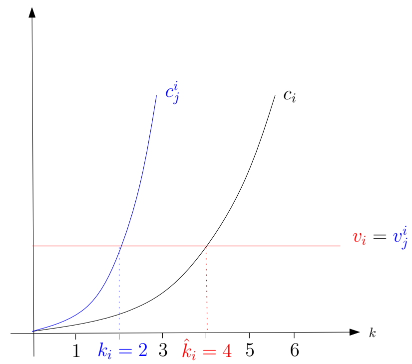

Figure 4: An example in which the cost c_i and value v_i are such that i's cognitive bound $\hat{k}_i = (c_i, v_i) = 4$. The behavioral level k_i is equal to 2 or 4 depending on whether the opponent is believed to be *less* or *more sophisticated* (respectively, on the left and on the right). In this example, for illustrative purposes we set $v_i = v_j^i$ and constant in k. Note that, in this case, if i believes that i is more sophisticated (i.e., $c_j^i(k) < c_i(k)$ for all k, as in the right panel), then the cognitive bound is binding ($\hat{k}_i = k_i = 4$).

The minimum operator here represents the idea that i's beliefs over j's steps of reasoning are bounded by his own cognitive bound, $\hat{k}_i$ (see Fig. 4). Player i then plays $a_i = a_i^{i,\hat{k}_i}$.

Note that this implies that a player always responds to either the opponent's action associated with the step where he thinks the opponent has stopped, or at the player's own maximum cognitive bound: in the latter case, the cognitive bound is *binding* in the sense that the player's beliefs about the number of steps undertaken by his opponent are limited by his own cognitive bound.

2.4 Endogenous Coordination in the BoS game

Consider player 1's value of reasoning in the baseline BoS game of Figure 3. When $a_1^{1,k-1} = B_1$, then $v_1(k) = \max\{r - r, 1 - 0\} = 1$, and when $a_1^{1,k-1} = W_1$, then $v_1(k) = \max\{r - r, r - 0\} = r$. Note that there is an asymmetry between the two actions: if, at step $k-1$, the player believes that B_1 is best, then the maximum gain he could obtain is 1; But if he believes that W_1 is best, then he has more to gain, and his value is now r. Furthermore, if r increases, the maximum gain at step in which $a_1^{k-1} = B_1$ is not affected and remains at 1, while it increases at steps in which $a_1^{1,k-1} = W_1$. Note also that this value of reasoning need not coincide with what the player will actually learn. For instance, whether the path of reasoning contains $a_1^{1,k-1} = a_1^{1,k} = B_1$ or, alternatively, $a_1^{1,k-1} = B_1$ and $a_1^{1,k} = W_1$, the value of reasoning for the k-th step is the same, and equal to 1. This is because the agent does not know what he will learn beforehand, otherwise it would imply that he has already performed the k-th step of reasoning (cf. Alaoui and Penta (2022)).

Observe that since the cost of reasoning increases unboundedly and the value function does not, for any r in the BoS game, for any player i and for his associated path of reasoning, there is a $\hat{k}_i(r)$ for which $c_i(\hat{k}_i) > v(\hat{k}_i)$, which is the stopping rule for player i at that r. This simple structure yields very sharp implications for any path of reasoning

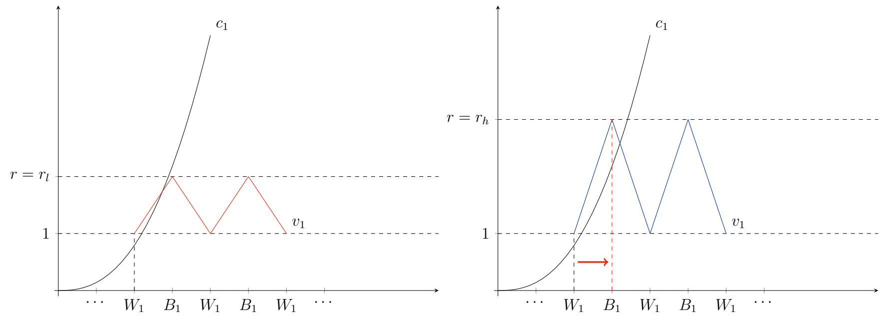

Figure 5: Low-payoff cognitive bound such that $a_1^{1,\hat{k}_1} = W_1$.

that is *responsive*.[5] For any such path of reasoning, and for any r, consider any player i with a responsive path, and whose last step of reasoning is $\hat{k}_i(r)$. Clearly, we have either $a_i^{i,\hat{k}_i} = B_i$ or $a_i^{i,\hat{k}_i} = W_i$. Suppose first that $a_i^{i,\hat{k}_i} = B_i$. Then $v_i(\hat{k}_i + 1) = 1$, and since the agent doesn't perform the $(\hat{k}_i + 1)$-th step, it must be that $c_i(\hat{k}_i + 1) > 1$. In this case, an increase in r has no effect on $v_i(\hat{k}_i + 1)$, and so the threshold $\hat{k}_i(r)$ remains unchanged as r goes up. Now suppose instead that $a_i^{i,\hat{k}_i} = W_i$. Then, $v_i(\hat{k}_i + 1) = r$ and $c_i(\hat{k}_i + 1) > r$. Since $c_i(\hat{k}_i + 1)$ is not infinite, there exists a finite r' such that $r' > c_i(\hat{k}_i + 1)$, given which the agent would perform at least one extra step. Take now the minimum $\tilde{k}_i \geq \hat{k}_i + 1$ for which $a_i^{i,\tilde{k}_i} = B_i$. Such a $\tilde{k}_i$ is guaranteed to exist, by the assumption that player i's path is responsive. For high enough r', this step will be reached, by the same argument as above. But at that step, it must be that the agent stops: he would only have continued if $1 \geq c_i(\tilde{k}_i + 1)$, but we know that $c_i(\tilde{k}_i + 1) > c_i(\hat{k}_i + 1) > r > 1$. Hence, here as well, player i's reasoning stops at B_i for a responsive path. This logic implies the following result:

Lemma 1 *Under the maintained assumptions on the cost and value of reasoning, for any $c_i(\cdot)$ and for any* responsive *path of reasoning, in the BoS game above there exists $\bar{r}_i$ such that, for all $r > \bar{r}_i$, player i stops reasoning at some step $\hat{k}(r)$ such that $a_i^{i,\hat{k}_i(r)} = B_i$.*

The logic of this result is illustrated in Figures 5 and 6, in which the (responsive) path of reasoning is such that $a_i^{i,k}$ alternates between B_i and W_i. This would be the case, for instance, for the level-k reasoning example provided previously, when the anchor is either (B_1, B_2) or (W_1, W_2), so that the path alternates between (B_1, B_2) and (W_1, W_2). It would also be the case for the deliberation over equilibria form of reasoning, if the player alternates between the two equilibria, (B_1, W_1) and (B_2, W_2). As can be seen in Figure 5, if 1's depth $\hat{k}_i$ for $r = r_l$ has associated $a_1^{1,\hat{k}_i} = W_1$, then a large enough increase in r (from r_l to r_h, in the figures) will lead to B_1. If, as in Figure 6, 1's depth $\hat{k}_1$ for lower r has associated $a_1^{1,\hat{k}_1} = B_1$, then an increase in r has no effect. Whereas the actual step $\hat{k}_1$

[5] As we explained earlier, if the path of reasoning is absorbing then reasoning has ultimately no impact on what is learned past the threshold at which the path stops changing. If the threshold of absorption $\bar{k}$ is greater than 0, then reasoning would have an effect until that threshold is met, but this will be essentially identical to considering the responsive case, and so we omit this discussion.

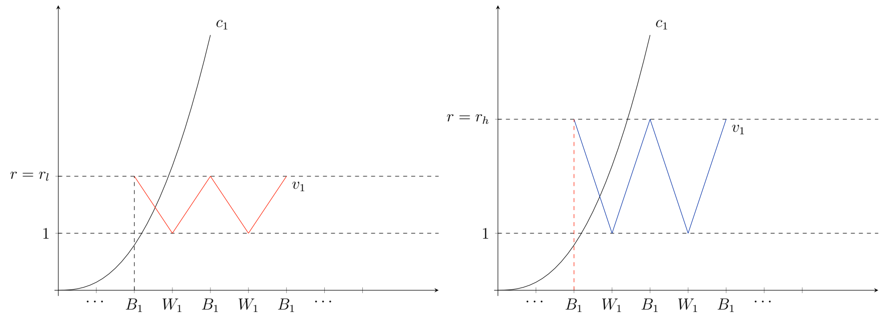

Figure 6: Low-payoff cognitive bound such that $a_1^{1,\hat{k}_1} = B_1$.

at which the agent stops may vary in the two cases, in either case it would be such that $a_i^{\hat{k}_1} = B_1$ for high enough r.

Note that applying the same logic as Lemma 1 to i's reasoning about j – i.e., using the cost and values c_j^i and v_j^i – yields the following implications for i's expectation of his opponent's depth of reasoning, $\hat{k}_j^i$:

Lemma 2 *Under the maintained assumptions, for any $c_j^i(\cdot)$ and for any* responsive *path of reasoning, in the BoS game above there exists $\bar{r}_j^i$ such that, for all $r > \bar{r}_j^i$, player i thinks that j stops reasoning at some step $\hat{k}_j^i(r)$ such that $a_j^{i,\hat{k}_j^i(r)} = B_j$.*

As noted in Remark 1, if a player thinks that the opponent is more (resp., less) sophisticated than he is himself – i.e., if $c_j^i \in C^+(c_i)$ (resp, if $c_j^i \in C^-(c_i)$) – then it implies that, with symmetric incentives to reason, he would expect his depth of reasoning to be weakly higher (resp., lower) than his own. In that remark, the inequality is *weak* because it may be that the cost functions are very close to each other, and hence for some levels of the value of reasoning they would effectively entail the same depth. The next assumption rules out this possibility, in that it requires that players' beliefs about the opponent's sophistication is different from one's own, in the sense that beliefs c_j^i are sufficiently lower (resp., higher) than c_i to effectively entail different depths of reasoning.

Formally: fix player i's path of reasoning in the BoS game, and type $t_i = (c_i, c_j^i)$. We say that i thinks that j is **strictly more (resp. less) sophisticated** than i if $c_j^i \in C^+(c_i)$ and if for every $r \geq 1$, $\mathcal{K}(v_i, c_i) < \mathcal{K}(v_i^j, c_i^j)$ (resp., $c_j^i \in C^-(c_i)$ and $\mathcal{K}(v_i, c_i) > \mathcal{K}(v_i^j, c_i^j)$). Given this, if i believes that j is more sophisticated, then i plays the action associated with i's cognitive bound, $\mathcal{K}(v_i, c_i)$, which by Lemma 1 induces action B_i for high enough r. If instead i believes that j is less sophisticated, then i thinks that j plays the action associated with *his* cognitive bound, that is B_j, and best-responds to that by choosing W_i. The next result follows:

Proposition 1 (Individual behavior in the BoS: Heterogeneous Sophistication) *Under the maintained assumptions, in the BoS game, for any* responsive *path of reasoning* *there exists $\bar{r}_i$ such that, for all $r > \bar{r}_i$, player i plays B_i if he thinks that j is* strictly more *sophisticated, and W_i if he thinks that j is* strictly less *sophisticated.*

Applying Proposition 1 to both players, delivers the following result:

Proposition 2 (Eductive Coordination in the BoS) *Under the maintained assumptions, in the BoS game, if both players' paths of reasoning are* responsive *and if they agree that i is strictly more sophisticated than j, there exists $\bar{r}$ such that, for all $r > \bar{r}$, players play $a = (W_i, B_j)$, the Nash equilibrium most favorable to player j.*

Proposition 2 provides our main result concerning how coordination can occur endogenously in the BoS game, when players believe they have different sophistication, and they agree about their relative ranking.

A second question concerns the case in which players do not believe that they have different sophistication. If a player thinks that his opponent is **equally sophisticated**, then within our model this means that he has equal costs. Formally, $c_i(k) = c_j^i(k)$ for all k. But, if that's the case, then players i chooses according to his own bound when facing j, and not according to his beliefs. This leads to the following result:

Proposition 3 (Individual behavior in the BoS: Equal Sophistication) *Under the maintained assumptions, in the BoS game, if i's path of reasoning is* responsive *and if i thinks that j is* equally sophisticated*, there exists $\bar{r}_i$ such that, for all $r > \bar{r}_i$, i plays B_i.*

Hence, note that player i behaves in the same way if he thinks that the opponent is 'equally' or 'more' sophisticated than he is himself: in both cases, his choice is driven by his own cognitive bound. Note as well that Proposition 3 implies that, if both players think that they are equally sophisticated, and if r is high enough, then they would end up playing an action profile that is not a Nash equilibrium:

Proposition 4 (Miscoordination in the BoS) *Under the maintained assumptions, in the BoS, if both players' paths of reasoning are* responsive *and if they agree that they are* equally sophisticated*, then there exists $\bar{r}_i$ such that players play $a = (B_i, B_j)$ for all $r > \bar{r}_i$.*

2.5 Focality, Alignment and Eductive Coordination: Discussion

Since Schelling (1960), a *focal point* is an action profile that is *salient*, *self-enforcing* (i.e., consistent with players' rationality), and such that players are firm in their expectation that it would occur. Hence, if such a focal point exists – be it due to payoff considerations (e.g., if efficiency, risk-dominance, etc., are shared refinement criteria), to 'non-mathematical' properties of the game (e.g., intrinsic characteristics or labeling of the actions, as in Crawford et al. (2008), Charness and Sontuoso (2022), etc.), to players' mode of cognition (e.g., Bilancini et al. (2017)), or to previous experience of play – then it is natural to expect agents to play accordingly. All these cases can be naturally mapped to our model as follows:

Definition 2 (Focal Points) *Profile a^* is* (subjectively) focal for player *i if it is a Nash equilibrium and $a^{i,k} = a^*$ for all k. Profile a^* is* focal *if it is focal for both players.*

In words: players start out with a common self-enforcing profile in mind (a Nash equilibrium), and further introspection confirms that it should be played.

Clearly, if players share a focal point, then equilibrium coordination is not an issue: the coordination problem is basically assumed away, and its explanation boils down to a *theory of focal points* (e.g., Sugden (1995)). The focus of our analysis instead is on whether coordination can be achieved *in the absence of a focal point.* Absence of a focal point may be due to two possibilities: (i) at least one of the players, subjectively, has no focal point; (ii) both players believe in a focal point, but not in the same. The second case may seem odd, but it's important nonetheless. For instance, within a level-k model of reasoning, a practical example would be that of an American and a British car driver, who come from opposite directions, and play the obvious coordination game in which they simultaneously choose whether to drive on the left or on the right. If not aware of the nationality of the opponent, they would (arguably) each embrace a social norm which is subjectively focal, but not shared. The miscoordination which would obviously arise in this case can be ascribed to the failure to recognize that the 'old' social norm does not apply to this particular situation, an instance of case (ii) above. In this thought experiment, it is natural to hypothesize that if the two drivers were made commonly aware of the nationality of the opponent, then the subjective $a^{i,1}$ would *not* be a NE, and hence players would not believe in any particular point being focal. Clearly, miscoordination would be possible in this situation, and it would instead be an instance of case (i) above.

Our model of reasoning implies that, while coordination could not be reached in the first example (the two drivers are not aware of the opponent's nationality, and hence their path of reasoning is *absorbing*), in the second case coordination can be achieved, despite the absence of a focal social norm, if two conditions are met (cf. Proposition 2): (i) first, players' payoffs display a sufficiently strong bias in favor of the 'own side' of the road (so that the game looks like a BoS, and the r is sufficiently high); (ii) second, if both players agree on their relative sophistication – that is, if they commonly believe (even if wrongly so) that player i is more sophisticated than player j. The fact that this result obtains as r grows unboundedly is perhaps counterintuitive, as one might expect that, at least for very high r, players would optimally switch to their favorite action. This intuition can be formalized by letting players entertain non-degenerate beliefs in their reasoning process: in this case, if they always attach a positive probability to the opponent conceding, then the condition that a_i^k is a best-response to player i's conjectures at step k implies that there exists a threshold $\hat{r}_i$ beyond which only the 'own' favorite action is considered. The model can clearly be extended in this direction, at the cost of less stark predictions (e.g., there would exist a parameter range, $(\bar{r}_i, \hat{r}_i) \subseteq \mathbb{R}$, within which the above coordination result obtains, but not if $r > \hat{r}_i$). Abstracting from the possibility of non-degenerate conjectures, as we do above, distills the essence of the logic at the heart of our coordination mechanism and delivers sharp and falsifiable predictions.

Similarly, the model can also be extended to account for non-degenerate beliefs about the opponent's cost of reasoning, as Alaoui and Penta (2016) do in the context of level-k reasoning. The propositions above would not be affected by such an extension. The reason is that the details of players' beliefs about the costs of reasoning do affect the exact position of the critical threshold $\hat{r}_i$, but not its existence. The fact that the results above obtain under very minimal restrictions on players' beliefs about each other is an important strength of the model, particularly from the viewpoint of its testability.

3 The Experiment

3.1 Experimental design and logistics

The experiment is designed to test whether behavior in the BoS game is in line with our predictions. It also includes other games, which will be discussed in the next section, that allow us to test whether the results can be explained by alternative theories instead.

At the beginning of the experiment, all subjects complete a *test of cognitive sophistication*. The test contains the Muddy Faces game (cf. Weber (2001)), a version of the Mastermind game and a centipede game. The questions are the same as in Alaoui and Penta (2016) and in Alaoui, Janezic, and Penta (2020). As a robustness check, around a third of subjects complete the Raven's Advanced Progressive Matrices (APM) test (Raven (1994)) rather than our test. To assess whether both tests can be used interchangeably, subjects complete the *alternative test* at the end of the experiment (subjects who first completed our test saw the APM test at the end of the experiment, and viceversa). Results for the comparison between the two tests are given in Appendix A.3.2.

We then separate subjects into three groups, depending on whether their scores where High, Moderate, or Low. The High and Low groups are informed of their labels, the Moderate group is not.[6] For the main experiment, we use only the High and Low groups, so as to obtain enough perceived distance in sophistication between the groups. This reflects the theoretical notion of heterogeneous sophistication underlying Propositions 1 and 2, which requires that the perceived difference in sophistication leads to effectively different depths of reasoning (p. 13). The Moderate group plays an unlabeled treatment. For the purpose of testing our model's predictions, only the High and Low groups are relevant, and all the discussions below refer to these groups. We document the behavior of the Moderate group in Appendix A.3.3.

The first game that subjects play is the following BoS game, which subjects play both against an opponent with the *same* label and against one with the *other* label:

[6]These three categories were defined by pre-determined cutoff scores, which were based on the distributions of test scores obtained in Alaoui and Penta (2016), Alaoui, Janezic, and Penta (2020) and subsequent pilots. Cutoffs are not determined session by session, because subjects of the entire sample played against one another, and were paid once all the sessions ended.

	W	Z
X	$r, 50$	$0, 0$
Y	$0, 0$	$50, r$

Figure 7: Battle of the Sexes

where $r \in \{51, 70\}$, depending on the treatment. The action labels (X, Y, W and Z) were chosen to avoid salience. For ease of mapping with the theoretical results in the previous section, below we will use B_i and W_i to refer to actions X (Z) and Y (W), respectively. Each subject played the following four versions of the game, in the role of the *row player*, without feedback and with random anonymous matching at every round:

- **BoS-Same-S:** The BoS game is played against someone with the *same* label, for the *smaller* reward $r = 51$.
- **BoS-Other-S:** The BoS game is played against someone with the *other* label, for the *smaller* reward $r = 51$.
- **BoS-Same-G:** The BoS game is played against someone with the *same* label, for the *greater* reward $r = 70$.
- **BoS-Other-G:** The BoS game is played against someone with the *other* label, for the *greater* reward $r = 70$.

In addition to the BoS game, subjects also played a version of the *modified* BoS game, which we will introduce below, as well as Stag Hunt and an Asymmetric Matching Pennies game. These games are included to assess the viability of some alternative mechanisms that may guide subjects' choices in these game, and will be discussed in Sections 4 and 5. Subjects in the experiment are matched randomly for each interaction, they are paid randomly for one version, out of four, of each game and they receive no feedback.

At the end of the experiment, subjects were asked whether they believe that performance in the initial test is correlated with success in the games. They then completed a short cognitive reflection test (CRT, Frederick (2005)), a hypothetical acyclical 11-20 game (Alaoui and Penta (2016)) and the alternative test of cognitive sophistication.

The experiments were conducted in Spring 2022 at the BES lab at Universitat Pompeu Fabra. It was coded using z-Tree (Fischbacher (2007)). In total, 183 subjects participated in the full experiment, spread over 16 sessions. They received an average pay of €21.5, including a €5 show-up fee, for an approximate duration of 110 minutes. Subjects were paid for one version of each game. Specifically, one out of the four versions was picked at random and this was repeated for each of the four types of games for the labeled treatment and one out of two for the unlabeled treatment. Of the 183 subjects, 149 participated in the labeled treatments, 43 of which were classified as Low and 106 as High, while 34 subjects were in Moderate group and participated to the unlabeled treatment.

3.2 Predictions for the BoS game

Let L denote the label of subjects who were classified as Low on the test, and H that of those who were classified as High. Recall that subjects are informed of these labels. The propositions from Section 2 then map to testable predictions for the experiment, under the following assumptions:

Assumption 2 (Identification Assumptions)

1. *Subjects of the same label commonly agree that they are* equally sophisticated.

2. *Subjects of different labels commonly agree that label H subjects are* strictly more sophisticated *than label L subjects.*

3. *Paths of reasoning are* responsive *for at least some percentage of L and H subjects.*

Assumptions 2.1 and 2.2 are the key assumptions for our exercise, and the entire experiment (particularly the way that labels were created and assigned) was designed in order to ensure that they are satisfied. They can of course be weakened to allow for some noise, but we keep them as they are for ease of exposition. We discuss possible variations in footnote 7, when we provide the predictions. As we explain below, strictly speaking Assumption 2.3 is not required for the predictions that follow. Without this assumption, however, our model has no bite, since all our predictions are predicated on the assumption that paths of reasoning are responsive. That being said, we stress that these are very weak assumptions, particularly within the context of our experimental design.

For $g \in \{L, H\}$, let $p^g(\cdot)$ denote the percentage of subjects in group g that play their own preferred action B_i, where the argument of the function refers to the treatment. The predictions that we provide follow directly from Assumption 2 and Propositions 1 and 4.

We first compare subjects' behavior as they play against someone with the *same* or with the *other* label, both for the smaller and for the greater reward (respectively, $r = 51$ and $r = 70$ – recall that these are denoted by S and G in the treatments):

Prediction 1 (Own to other label comparison)

1. $p^H(\textit{BoS-Same-S}) \geq p^H(\textit{BoS-Other-S})$ *and* $p^H(\textit{BoS-Same-G}) \geq p^H(\textit{BoS-Other-G})$*: the percentage of High subjects playing their own preferred action in the BoS game is* lower *when playing against subjects with the other label than against subjects with the same label, for both values of r.*

2. $p^L(\textit{BoS-Same-S}) = p^L(\textit{BoS-Other-S})$ *and* $p^L(\textit{BoS-Same-G}) = p^L(\textit{BoS-Other-G})$*: the percentage of Low subjects playing their own preferred action in the BoS game is* the same *when playing against subjects with the other label than against subjects with the same label, for both values of r.*

Note that while the predictions in point 1 above involve weak inequalities, those in point 2 are in terms of equalities. Retracing the logic from Section 2.4, the difference is due to the following: Label H subjects play according to their beliefs about label L subjects (which are viewed to be *less sophisticated*) but according to their own cognitive bound when they play against label H subjects. This can lead to a difference in behavior, and shift some subjects to the opponent's preferred action when they play against an L subject. Label L player, however, play according to their bound both against L (viewed as *equally sophisticated*) and against H (viewed as *more sophisticated*), and therefore their behavior does not change.[7] Note as well that these predictions also holds without Assumption 2.3, because if no subjects have a responsive path then their behavior would not change, which is allowed by the weak inequalities in Prediction 1. But if that were the case, then the mechanism discussed here would never be switched on, and hence we would observe no change even as payoffs are varied, which we turn to next. As we will discuss when we present our findings, the experimental results are indeed consistent with Assumption 2.3.

In addition to the predictions above, in which we consider behavior as the opponent's label is varied but the payoffs remain the same (for both values of r), we also make the next set of predictions as r is varied, but the opponent is kept fixed.

Prediction 2 (Low to high payoffs comparison)

1. $p^H(\text{BoS-Same-S}) \leq p^H(\text{BoS-Same-G})$ *and* $p^H(\text{BoS-Other-S}) \geq p^H(\text{BoS-Other-G})$*: the percentage of high subjects playing their own preferred action is weakly* increasing *in r when playing their own label in the BoS game, and weakly* decreasing *when playing the other label.*

2. $p^L(\text{BoS-Same-S}) \leq p^L(\text{BoS-Same-G})$ *and* $p^L(\text{BoS-Other-S}) \leq p^L(\text{BoS-Other-G})$*: the percentage of low subjects playing their own preferred action in the BoS game is weakly* increasing *in r both when playing their own label and the other label.*

The predictions above, which also follow directly from Assumption 2 and Propositions 1 and 4, show some more subtle implications of our model which would arguably be challenging to replicate with other mechanisms or alternative explanations. The reasoning behind these predictions is as follows. In Prediction 2.1 above, the high type is playing according to his bound when playing his own label, and as r increases, there is a higher percentage of high types whose bound will be at their own preferred action (as demonstrated in Section 2.4). In particular, subjects for whom $51 < \bar{r}_i$ while $70 > \bar{r}_i$, the action

[7] Assumption 2.1 could be weakened, for instance, by allowing that a majority of the subjects believes that those of the same label as themselves are equally (or more) sophisticated, while others believe that those of the same label are less sophisticated. In that case, Prediction 1.1 would remain unaffected, while Prediction 1.2 would change to weak inequalities rather than equalities. This is because those who believe others of the same label are less sophisticated would play according to their beliefs, and not their cognitive bound. This alternate assumption would therefore be more permissive in what it allows from our results in testing the theory. As previously discussed, however, we maintain the simpler Assumption 2.1, which requires a more demanding test of our theory (and analogously if we were to weaken Assumption 2.2).

they choose may switch as their cognitive bound is increased when r goes up, while for other subjects (those for whom $51 > \bar{r}_i$) their action at the cognitive bound may already be at B_i for $r = 51$. For this latter group, increasing r has no effect. When playing against the low label, however, his bound is irrelevant. Rather, a higher percentage of H label believe that their opponent's bound will stop at their (L's) preferred action, and reacts accordingly. Label L players, instead, play according to their bound whether their opponent is L or H. As r increases, there is a percentage of subjects for whom this bound may switch to their own preferred action.

Finally, we provide the following predictions, which relate the strength of the belief-effect with the change in the stakes of the game:

Prediction 3 (Beliefs-payoffs interaction)

1. $p^H(\textit{BoS-Same-G}) - p^H(\textit{BoS-Other-G}) \geq p^H(\textit{BoS-Same-S}) - p^H(\textit{BoS-Other-S})$: *For the high label subjects, the belief effect is weakly stronger in the high-payoff game.*

2. $p^L(\textit{BoS-Same-G}) - p^L(\textit{BoS-Other-G}) = p^L(\textit{BoS-Same-S}) - p^L(\textit{BoS-Other-S})$: *For the low label subjects, the belief effect is the same in the two payoff treatments.*

While all of the predictions above are about individual behavior, when discussing our findings below, we will also discuss whether there is increased coordination in the heterogeneous treatments on the preferred action profile of the low label subjects. Formal predictions on coordination comparing homogeneous to heterogeneous treatments for equal payoffs would require stronger assumptions on the comparability of the L and H groups, which we have not imposed.

3.3 Results of the BoS Game

In this section, we discuss the results relating to Predictions 1 and 2, which are tested using the BoS game. Recall that subjects always choose in the role of the row player.

According to Prediction 1.1, High label subjects are more likely to play their own preferred action in the BOS game (in this case X) against High label players than against Low label players, both for low and high payoffs. Table 1 shows the percentage with which High label players choose their preferred action X against each opponent type and for both payoff versions of the BoS game.

Analyzing first the low payoff version of the game, we observe that around 54% of High label players choose their preferred action X when they play against another player with the High label (see Table 1). However, when they face a Low label player, this percentage drops to 34.91%. We conduct a panel regression with the High players, restricting the sample to the low payoff version of the BoS game, and find that the coefficient on a dummy of whether they are playing against a High or Low label opponent is significant at the 1% level (see Table 2 Models (1) and (2)). We also compare the distribution of chosen

H players	H opponent	L opponent	% pt difference across opponents
Low payoff	54.37%	34.91%	19.46
High payoff	67.92%	36.79%	31.13
% pt difference across payoffs	13.25	1.88	

Table 1: Results BoS Game - High Label Players: % choosing X (their preferred action). Note that percentage point differences are given in absolute terms.

actions using a Wilcoxon signed rank test. The p-value of the test statistic is 0.002. These findings are all consistent with Prediction 1.1.

Repeating the analysis for the high payoff version of the game, we find that 67.92% play X against a High opponent but that only 36.79% play X against a Low opponent. The regression coefficient is significant at more than 0.1% (see Table 2 Models (3) and (4)). The p-value of the Wilcoxon signed rank test statistic is also less than 0.001. This shows that for both payoff versions of the BoS game, High label players play their preferred action, X, significantly less when they play against a Low label opponent than against a High label opponent. This lends further support to Prediction 1.1.

	(1) Choice of X	(2) Choice of X	(3) Choice of X	(4) Choice of X
Opponent has H label	0.179***	0.179***	0.311***	0.311***
	(2.79)	(2.79)	(4.60)	(4.59)
High Confidence Test		-0.0106		0.0287
		(-0.09)		(0.35)
Constant	0.349***	0.358***	0.368***	0.344***
	(10.88)	(3.25)	(10.87)	(4.18)
Observations	212	212	212	212

t statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Table 2: BoS Game: Panel regression results for H label players
Models (1) and (2) give the results for the low payoff versions of the BoS game and (3) and (4) for the high payoff versions. Standard errors are clustered at the subject level.

We next turn to Prediction 1.2, which predicts that the Low label subjects would be as likely to play their own preferred action against Low label as against High label opponents, for both low and high payoffs. The results for the Low label group are displayed in Table 3. In the low payoff BoS game, we find that 53.49% of Low subjects play their preferred action, X, against a Low label opponent. This percentage increases to 62.79% when playing against an opponent from the High label group. This difference is not statistically significant (p-value= 0.421). We run equivalent panel regressions to the H label discussed above. The regression results in Table 4 Models (1) and (2) show that there is no significant effect of playing against a high label opponent. The p-value of the Wilcoxon signed rank test statistic is 0.414. The lack of significance in the regressions is

L players	L opponent	H opponent	% pt difference across opponents
Low payoff	53.49%	62.79%	9.3
High payoff	60.47%	60.47%	0
% pt difference across payoffs	6.98	2.32	

Table 3: Results BoS Game - Low Label Players: % choosing X (their preferred action). Note that percentage point differences are given in absolute terms.

consistent with Prediction 1.2.

For the high payoff version of the game, we find that 60.47% of Low label players choose their preferred action, irrespective of the label of their opponents. The regression coefficient is not significant (Table 4 Models (3) and (4)) and the Wilcoxon signed rank test statistic is also not significant (all have p-value= 1). This is again consistent with Prediction 1.2. Testing next Prediction 2.1, we first consider whether it is the case that

	(1)	(2)	(3)	(4)
	Choice of X	Choice of X	Choice of X	Choice of X
Opponent has H label	0.0930	0.0930	0	1.46e-16
	(0.81)	(0.80)	(0.00)	(0.00)
High Confidence Test		0.113		-0.0281
		(1.13)		(-0.27)
Constant	0.535***	0.477***	0.605***	0.619***
	(9.30)	(5.11)	(10.89)	(6.65)
Observations	86	86	86	86

t statistics in parentheses

* p<0.10, ** p<0.05, *** p0.01

Table 4: BoS Game: Panel regression results for L label players
Models (1) and (2) give the results for the low payoff versions of the BoS game and (3) and (4) for the high payoff versions. Standard errors are clustered at the subject level.

High label subjects are (weakly) *more* likely to play their own preferred action against High labels as payoffs are increased, and *less* likely to play their preferred action when playing Low, as payoffs are increased. We indeed see from Table 1 that when High label subjects play against High opponents, 54.37% play X for the low payoff treatment, compared to nearly 68% for the high payoff treatment. This difference is statistically significant as shown in the corresponding panel regression results given in Model (1) of Table 15 in the Appendix. These results therefore support Prediction 2.1. When playing against Low opponents, the percentage playing X increases from roughly 35% to around 37%, but this is not significant.

Prediction 2.2 states that Low label subjects are (weakly) less likely to play their preferred action for low payoffs compared to high payoffs, against both Low and High labels. We see from Table 3 that the percentage of Low label subjects playing X does increase from 53.49% to 60.47% against Low opponents, which is consistent with the prediction. The percentage decreases against high from 62.79% to 60.47%, but these results are not significant (see Table 16 in the Appendix).

We now test Prediction 3.1, which states that for the High label subjects, the belief effect is (weakly) stronger in the high payoff game. The percentage point difference (last column of Table 1) increases from 19.46% for the low payoff version to 31.13, which is not statistically significant (Wilcoxon signed-rank test, p-value =0.264). For Prediction 3.2, which predicts no difference in the belief effect across the payoffs, we see a decrease from 9.3 to 0 (Table 3). This effect is not statistically significant (Wilcoxon signed-rank test, p-value =0.519).

Jointly, therefore, the findings are consistent with our predictions for the BoS game. Moreover, while Predictions 1.1 and 2 are in terms of *weak* inequalities, the fact that we observe that a percentage of subjects change behavior is indicative that the payoffs used in the experiment are sufficiently high for our model to have bite, and indeed the predictions are confirmed with strict inequalaties.

While our predictions are for individual behavior, we close this section by analyzing whether coordination is more likely to occur under heterogeneous treatments (High vs. Low labels) than homogeneous treatments (High vs High and Low vs Low), and on the preferred action profile of the Low labels. When we examine coordination outcomes for the low payoff BoS where L and H players are matched with each other, we find that 40.87% coordinate on the equilibrium most favorable to the L players (Tables 5 and 6).

H row player: percentage of (Y,Z)	Low payoff	High payoff
vs. H opponent	18.86%	23.71%
vs. L opponent	40.87%	38.22%

Table 5: Results BoS Game - H label row players, % coordination on the opponents' favorite equilibrium.

L row player: percentage of (X,W)	Low payoff	High payoff
vs. L opponent	28.00%	34.00%
vs. H opponent	40.87%	38.22%

Table 6: Results BoS Game - L label row players, % coordination on the player's favorite equilibrium.

When we compare this to the frequency with which the same equilibria are achieved when subjects are matched with an opponent of the same label, it becomes apparent how strongly this increases under heterogeneous matching.[8] When H play against other H, we find that only 18.86% coordinate on the (Y, Z) equilibrium under homogeneous matching, which corresponds to the action profile in which the row players play the equilibrium most favorable to their opponent. Thus, this percentage is less than half that under

[8]To calculate the coordination percentages for the homogeneous treatments, we split the groups according to their exogenous row - column classification. For the heterogeneous treatments instead, Tables 5 and 6 provide the combined percentages, given the interchangeability of the two groups.

heterogeneous matching. In fact, even if we consider, for H versus H, the equilibrium most favorable to the row players (X, W), we find that the percentage of coordination is 31.56%, which is also lower than the 40.87% who coordinate on the equilibrium favorable to the L players in the heterogeneous treatment. Similarly, considering L vs L, when we consider (X, W) the percentage of coordination is 28.00%, more than twelve points lower than 40.87%. Even if we consider (Y, Z), the equilibrium favorable to the column players, coordination is equal to 22.00%, which is even lower.

In the case of the high payoff version, in the heterogeneous treatment 38.22% coordinate on the equilibrium favorable to the L players. In the homogeneous treatment with H vs H, we find that 23.71% coordinate on the (Y, Z) equilibrium, which again is substantially lower than for the heterogeneous treatment. Again, even if we consider (X, W) instead (for H vs H), we find that 19.94% coordinate, which is markedly lower than 38.22%. For the L vs L subjects, 34.00% coordinate on (X, W), which is around four points lower than 38.22%, and 16.00% coordinate on (Y, Z), which is less than half of 38.22%. Again, heterogeneous matching leads to a marked increase in coordination on this equilibrium, compared to either H vs H or L vs L, and for either equilibrium profile of the homogeneous treatments. Overall, therefore, our results show that coordination on the equilibrium preferable to the L player increases considerably when matching is heterogeneous, both for low and high payoff.

4 Competing Explanations

There are at least two alternative theories that one might consider to explain the experimental results shown in the previous section. The first is the view that the increased coordination that we observed in the heterogeneous treatments in the main experiment is merely the result of the asymmetry in the group labels, which may themselves serve as a coordination device (*label focality*). This view, however, is inconsistent with our finding that H label subjects react to their opponent while L label subjects do not.

The second competing explanation is the view that our mechanism is akin to granting the low type a sort of *first-mover advantage*, in the sense that it is *as if* the low type "commits" to stop reasoning first, at his preferred action profile, while the high type then concedes. To see that this is not an adequate way to summarize the insights of our model, consider the *modified BoS* game that we presented in the introduction, which we reproduce in Fig.8 labeling players' actions B_i and W_i to denote, respectively, the action associated with the *'best'* and *'worst'* equilibrium for player i.

Now consider player 1's value of reasoning in this game: When $a_1^{1,k-1} = B_1$, then $v_1(k) = \max\{70 - 70, 50 - (-r)\} = 50 + r$, and when $a_1^{1,k-1} = W_1$, then $v_1(k) = \max\{50 - 50, 70 - 0\} = 70$. Note that, compared to the canonical BoS game discussed in Section 2, the role of the B_i and W_i actions is now reversed: in the *modified BoS*, for $r > 20$, it is

	W_2	B_2
B_1	$70, 50$	$-r, -r$
W_1	$0, 0$	$50, 70$

Figure 8: A *modified* BoS, with $r \geq 20$.

now the W_i action that is associated with the higher value of reasoning, and this value is increasing with r. Thus, the same argument that led to the result in Proposition 2 for the BoS delivers the following conclusion for the *modified BoS*:

Proposition 5 (Eductive Coordination in the *modified BoS*) *In the* modified BoS *game, if both players agree that i is strictly more sophisticated than j, and if their paths of reasoning are* responsive*, there exists $\bar{r}$ such that, for all $r > \bar{r}$, players play $a = (B_i, W_j)$.*

Hence, in the *modified BoS*, the prediction of our model is opposite to the view that the low type receives a *first-mover advantage*: in this game, such a theory would predict the equilibrium most favorable to the player who is regarded to be of *lower* strategic sophistication; our model delivers the opposite prediction. Similarly, the *label focality* argument predicts that coordination occurs on the equilibrium preferred by the same label in both games, and thus contrasts with the predictions of our model across the two games. Hence, the two games together can be used to discern between our model and these competing explanations.

4.1 Experimental Design and Theoretical Predictions

To test whether the first-mover advantage view better explains the data than our model, we use the following game, which we refer to as the *Reverse Strategic Advantage* (RevSA):

	B_2	W_2
B_1	$130, 130$	$230, r$
W_1	$r, 230$	$170, 170$

Figure 9: The *Reverse Strategic Advantage* (RevSA) Game, with $r \in [190, 220]$.

This game generates a strategic advantage similar to the *modified BoS*, and leads to analogous predictions within our model, which are distinct from those of the BoS game. To see this, first note that the lower bound on the r parameter ensures that both (B_1, W_2) and (W_1, B_2) are equilibria, whereas the upper bound ensures that the ranking over the two is maintained within the parameter range (that is, $u_i(B_i, W_j) > u_i(W_i, B_j)$). Second, the value of reasoning is such that, when $a_1^{1,k-1} = B_1$, then $v_1(k) = \max\{230 - 230, r - 130\} = r - 130$, and when $a_1^{1,k-1} = W_1$, then $v_1(k) = \max\{r - r, 230 - 170\} = 60$. Thus, similar to the *modified BoS*, within the relevant parameter range, $r \in [190, 220]$, it is again the W_i action that is associated with the higher value of reasoning, and such a value is increasing with r. The reason for using this game rather than the *modified BoS* from Fig. 8 is to avoid negative payoffs, and to make it clear to an inattentive subject going through the experiment that they haven't seen this game previously.

The game used in the experiment takes exactly the form given in Figure 9, but adopting the X, Y and W, Z labels for the actions of the row and column players, respectively, and letting $r \in \{190, 220\}$, depending on the treatment. As with the BoS, the subjects played four versions of this game: with $r = 190$ against someone with same label or against someone with the other label, and with $r = 220$ against an opponent with each label.

By the same logic above, and under the same identification assumptions discussed in Section 3.2, the model delivers the following predictions for the *RevSA* game:

Prediction 4 *In the RevSA game, the predictions of our model are as follows:*

1. p^H(*RevSA-Same-G*) $\leq p^H$(*RevSA-Other-G*)*: the percentage of High subjects playing their own preferred action in the RevSA game with sufficiently high values of r is* higher *when they play against subjects with the other label than against subjects with the same label.*

2. p^L(*RevSA-Same-G*) $= p^L$(*RevSA-Other-G*)*: the percentage of Low subjects playing their own preferred action in the RevSA game with sufficiently high values of r is* the same *when they play against subjects with the other or with the same label.*

Note that these predictions are opposite to the view that the low type gets a first-mover advantage (FMA) under heterogenous matching, which we summarize as follows, maintaining the same notation as in the predictions above:

Prediction FMA *In the RevSA game, the predictions of the FMA-view are the following:*

1. p^H(*RevSA-Same-G*) $> p^H$(*RevSA-Other-G*)

2. p^L(*RevSA-Same-G*) $< p^L$(*RevSA-Other-G*).

4.2 RevSA Game Results

Table 7 contains the choices of the High label players in the four versions of the RevSA game. For the High label group, close to 19% of players choose their preferred action, X, when playing against someone with a High label in the low payoff version of the RevSA game. When the opponent changes to being a Low label player, this increases to nearly 35%. We again conduct panel regressions (Table 17 in the Appendix) to assess whether changing the opponent has a significant effect on the action choice and find that this effect is significant at the 1% level (see Models (1) and (2)). As with the BoS game, we also conduct a Wilcoxon signed rank test to confirm whether the two distributions of actions are statistically significantly different when the opponent's label changes. We obtain a p-value of 0.0095. These results go against the first-mover advantage explanation in that the High label players choose their preferred action more frequently when playing against a Low label player in the RevSA game while the opposite is true for the BoS game.

In the high payoff version of the game, we find that nearly 23% choose their preferred action when playing against another High label player. This percentage increases to more than 25% when the opponent changes to being a Low label player. However, this increase is not statistically significant (p-value of 0.62 in the panel regressions, Models (3) and (4) in Table 17, and of 0.59 for the Wilcoxon signed rank test). Thus, also under high payoffs, the High label players do not concede to the Low label players. We can therefore reject the first-movers advantage argument.

H players	Against H label	Against L label	% point difference across opponents
Low payoff	19.23%	34.91%	15.58
High payoff	22.86%	25.47%	2.61

Table 7: Results RevSA - High Label Players: % choosing X (their preferred action)

For the Low label players, more than 41% choose their preferred action when playing against another Low label player while around 47% select their preferred action against a High label opponent (see Table 8). This difference is not statistically significant (p-value= 0.67 for Model (1) and p-value= 0.69 for Model (2) in the regression, Table 18, and p-value= 0.695 for the Wilcoxon signed rank test). In the high payoff version, just over 38% select their preferred action against a Low label opponent and close to 33% against a High label opponent. Again, this difference is not statistically significant (p-value= 0.586 for Model (3) and p-value= 0.579 for Model (4) of the panel regressions, Table 18 in the Appendix, and p-value= 0.617 for the Wilcoxon signed rank test). The finding that there are no statistically significant changes to the choices by the Low label players is consistent with the model's predictions.

L players	Against L label	Against H label	% point difference across opponents
Low payoff	41.46%	46.51%	5.05
High payoff	38.10%	32.56%	5.54

Table 8: Results RevSA - Low Label Players: % choosing X (their preferred action)

While we can use the RevSA game to assess whether the first-mover advantage is a likely explanation for observed behaviour, our model does not make a prediction for games in which the value of reasoning is flat, and hence it cannot make a prediction for the low payoff version of the RevSA game. For the high payoff version, we predict that a greater fraction of High label subjects chooses X against a Low, compared to a High, label opponent. As stated above, we find that under the high payoff version, behavior changes in the desired direction but that the change is small. Observed behavior is also consistent with the existence of beliefs over *noise players*. For instance, if a fraction p of players believe that the opponents play either W or Z with equal probability, then the fraction of row players who play Y should increase. This may explain why such a large fraction of players selects Y (as well as the increase in the average number of players who choose Y as payoffs increase). Notice that the existence of noise players would be consistent with our predictions and is consistent with what we observe.

5 Additional Games

In addition to the BoS and the RevSA games, subjects also played four versions of a Stag Hunt game and of an Asymmetric Matching Pennies. In this section we discuss our findings for these additional games.

5.1 Stag Hunt

Stag Hunt was included to assess whether risk dominance might provide an alternative explanation for observed behavior. This game, shown in Fig.10, also uses a standard set-up with a low and a high payoff version ($r = 50$ and $r = 70$ respectively). As with the BoS and the RevSA games, subjects played each of the two payoff versions against a Low and against a High opponent and were informed of their opponent's label.

	W	Z
X	r, r	$0, 30$
Y	$30, 0$	$30, 30$

Figure 10: The *Stag Hunt* Game, with $r \in \{50, 70\}$.

In general, we find that a large majority of subjects chooses X, which is the preferred action. For the High label players in the low payoff version of the game, we find that nearly 81% choose X when facing a High label opponent while around 75% select X against a Low label opponent. For the high payoff version, we find similar results in that nearly 86% select X against a High label opponent and close to 81% against a Low label opponent (see Table 9). The differences in the frequency with which X is played across opponent labels are not statistically significant (Wilcoxon signed rank test).

H players	Against H label	Against L label	% point difference across opponents
Low payoff	80.95%	75.47%	5.48
High payoff	85.85%	81.13%	4.72

Table 9: Results Stag Hunt - High Label Players: % choosing X (their preferred outcome)

For the Low label players in the low payoff version, we observe that close to 77% select X against both another Low label player or a High label player. For the high payoff version of the game, we find that around 88% select X against a Low label opponent, while close to 74% select X against a High label opponent (see Table 10). For the high payoff version, the difference in behavior across different opponent labels is statistically significant at the 10% level as measured by a Wilcoxon signed rank test (p-value =0.058).

L players	Against L label	Against H label	% point difference across opponents
Low payoff	76.74%	76.74%	0
High payoff	88.37%	74.42%	13.95

Table 10: Results Stag Hunt - Low Label Players: % choosing X (their preferred outcome)

These results are largely consistent with our predictions. Note that in this game we predict that all subjects are (weakly) more likely to choose X for a sufficiently high payoff version of the game, and that there should not be a change in likelihood of playing X against the Low type compared to the High.

Also note that the result that subjects are more likely to choose X for the high payoff version is also consistent with risk dominance, since Y is risk dominant for the low payoff version and X is dominant for the high payoff version. But what is perhaps more surprising is that a large majority of subjects chooses X even for the low payoff version, which goes against risk dominance.[9] This result is fully consistent with our model, however. The value function here is asymmetric, in that there is a larger gain from continuing reasoning at Y than at X, both for the low and high payoff versions. Therefore, while we fully expect that risk dominance would be the dominant force for lower payoffs, it is noteworthy that the mechanism described in this paper may overtake risk dominance close to the threshold at which the payoff switch should theoretically occur.

5.2 Asymmetric Matching Pennies

Subjects also played an Asymmetric Matching Pennies (AMP) game. As with the other games, they played four versions, a low payoff and a high payoff version against both a Low and a High label opponent. In the low payoff version, the incentives of the row player are nearly flat, i.e. the asymmetry is slight. The asymmetry is much more pronounced in the high payoff version. As such, we expect to find a larger effect on behavior for the high payoff version. From the perspective of our model, the opponent has a flat value function and so, without additional assumptions on the path of reasoning, we can only predict that as payoffs increase, the frequency with which X is chosen should increase. This is exactly what we observe for both labels.

	W	Z
X	$r, 20$	$20, 40$
Y	$20, 40$	$40, 20$

Figure 11: The *Asymmetric Matching Pennies* (AMP) Game, with $r \in \{41, 160\}$.

Considering first the row players, we find that 58.82% of the High label players select

[9]Assuming a concave utility for money would not explain this result either.

X against another High label player, while nearly 55% select X against a Low label player, in the low payoff version of the game (see Table 11). This difference is not statistically significant as measured by a Wilcoxon signed rank test (p-value =0.819). For the high payoff version of the game, 62.26% of the High label players select X against a H opponent and 60.38% against a L opponent (difference is not statistically significant as measured by a Wilcoxon signed rank test). For the Low label row players, we find that 52% play X against another Low label player (see Table 12). Against a High label player, this increases to 60%. For the high payoff specification, 64% of Low label row players select X against either label. Differences in behaviour across opponent labels are not statistically significant (Wilcoxon signed rank test).[10]

H row players	Against H label	Against L label	% point difference across opponents
Low payoff	58.82%	54.72%	4.10
High payoff	62.26%	60.38%	1.88

Table 11: Results AMP - High Label row Players: % choosing X (their preferred outcome)

L row players	Against L label	Against H label	% point difference across opponents
Low payoff	52.00%	60.00%	8.00
High payoff	64.00%	64.00%	0.00

Table 12: Results AMP - Low Label row Players: % choosing X (their preferred outcome)

Considering next the column players, we find that High label column players in the low payoff specification choose Z with 69.23% frequency against High label players. Against Low label players, they choose Z with close to 68%. This percentage increases to 75% and 79.25% against a High, resp. Low, label opponent when they play the high payoff version of the AMP game. The results are given in Table 13. The High label column players thus best respond to the row player selecting X and anticipate the increase in the choice of X after their opponent's payoff from playing X increases. For the Low label column player, we find that close to 47% choose Z against another Low label player in the low payoff specification. This percentage increases to nearly 56% against a High label opponent. For the high payoff specification, 72.22% choose Z against a Low label opponent and nearly 67% against a High label opponent. Table 14 gives the results. Differences in behavior across opponent labels is not statistically significant for either L or H label column players (examined using a Wilcoxon signed rank test). Note that the model implies that the column player's value function is flat, but their opponent's is not. For the low payoff version, the opponent's value function is nearly flat so that we do not have a clear prediction of which action column players should select. However, for the high payoff

[10] In comparison with Goeree and Holt (2001), results for the low payoff version are similar to their symmetric matching pennies game, while slightly more H label subjects choose X, likely owing to the small asymmetry. While the effect for the high payoff version, with the large asymmetry, is smaller than in their experiment, the effect goes in the same direction.

specification, the row players are predicted to choose X more frequently. If the column players form beliefs about their opponent's incentives, here, they should best respond by playing Z more frequently. Our findings are fully consistent with this prediction.

H column players	Against H label	Against L label	% point difference across opponents
Low payoff	69.23%	67.92%	1.31
High payoff	75.00%	79.25%	4.25

Table 13: Results AMP Game - High Label column Players: % choosing Z

L column players	Against L label	Against H label	% point difference across opponents
Low payoff	47.06%	55.56%	8.5
High payoff	72.22%	66.67%	5.55

Table 14: Results AMP Game - Low Label column Players: % choosing Z

6 Conclusion

Individuals often face coordination problems in which they interact infrequently with others, and must attempt to coordinate with very little information on their opponent purely based on reasoning. These individuals may, however, form beliefs over their relative cognitive sophistication. In such situations, whether coordination can be achieved at all has been an important open question in game theory. In this paper, we have proposed a novel mechanism for how coordination can be achieved. When the individuals agree on each other's relative sophistication, then our model makes predictions concerning individuals' choices as a function of their incentives and beliefs over their opponents, which in turn can lead to increased coordination when the players agree that they are of different sophistication. In other words, while it is common to view homogeneity and shared culture as leading to increased coordination by inducing focal points (e.g., Kets and Sandroni (2019), Kets et al. (2022), and Kets (2022)) in the absence of focal points, we show that it is *heterogeneity* that leads to coordination. Note that here, rather than agreeing on the norms, players agree on relative cognitive abilities. But this agreement is not in itself enough – if players believe that they have similar sophistication, then they are *less* likely to coordinate in canonical games such as the BoS. Our model further predicts that, in the case of the BoS game with heterogeneous sophistication, the increased coordination occurs on the preferred outcome of the less sophisticated player. This may perhaps seem surprising, under the view that the more sophisticated player should be the one to 'win'. When testing our joint predictions for the BoS game in an experiment, we find strong support for our model.

At the same time, our mechanism might seem reminiscent of a kind of first mover advantage (FMA), in which the player viewed to be less sophisticated has the advantage of stopping reasoning first, so that the more sophisticated one must concede. We show, however, that in different games, our model makes predictions distinct from the ones that would obtain under FMA. We test this alternative mechanism and find that it is inconsistent with the subjects' behavior. We also conduct experiments using well-known additional games (stag hunt and an asymmetric matching pennies game) and again find support for our model. Taken jointly, the experimental results strongly support the mechanism introduced in this paper.

In closing, we note that the role of cognitive sophistication, and specifically beliefs over relative cognitive sophistication, has increasingly been recognized in game theoretical settings (cf. Proto, Rustichini, and Sofianos (2019, 2021); Lambrecht, Proto, Rustichini, and Sofianos (2022)).[11] This paper shows that such beliefs can play an important role even in achieving coordination in isolated settings.

[11] We note that despite the important differences between the underlying strategic environments, the qualitative predictions that we obtain from our model are in line with the experimental findings of Proto, Rustichini, and Sofianos (2019, 2021) and Lambrecht, Proto, Rustichini, and Sofianos (2022), on the effects that these beliefs have on the behavior in the repeated BoS.

References

Alaoui, L., K. A. Janezic, and A. Penta (2020): "Reasoning about others' reasoning," *Journal of Economic Theory*, 189, 105091.

Alaoui, L. and A. Penta (2016): "Endogenous depth of reasoning," *The Review of Economic Studies*, 83, 1297–1333.

——— (2022): "Cost-benefit analysis in reasoning," *Journal of Political Economy*, 130, 881–925.

Alós-Ferrer, C. and J. Buckenmaier (2021): "Cognitive sophistication and deliberation times," *Experimental Economics*, 24, 558–592.

Bilancini, E., L. Boncinelli, L. Luini, et al. (2017): "Does focality depend on the mode of cognition? Experimental evidence on pure coordination games," Tech. rep., Department of Economics, University of Siena.

Binmore, K. (1987): "Modeling rational players: Part I," *Economics & Philosophy*, 3, 179–214.

——— (1988): "Modeling rational players: Part II," *Economics & Philosophy*, 4, 9–55.

Charness, G. and A. Sontuoso (2022): "The Doors of Perception: Theory and Evidence of Frame-Dependent Rationalizability," *American Economic Journal: Microeconomics*.

Crawford, V. P., M. A. Costa-Gomes, and N. Iriberri (2013): "Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications," *Journal of Economic Literature*, 51, 5–62.

Crawford, V. P., U. Gneezy, and Y. Rottenstreich (2008): "The power of focal points is limited: Even minute payoff asymmetry may yield large coordination failures," *American Economic Review*, 98, 1443–58.

Esteban-Casanelles, T. and D. Gonçalves (2020): "The effect of incentives on choices and beliefs in games: An experiment," .

Fe, E., D. Gill, and V. Prowse (2022): "Cognitive skills, strategic sophistication, and life outcomes," *Journal of Political Economy*, 130, 2643–2704.

Fischbacher, U. (2007): "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental economics*, 10, 171–178.

Frederick, S. (2005): "Cognitive reflection and decision making," *Journal of Economic perspectives*, 19, 25–42.

Gill, D. and V. L. Prowse (2022): "Strategic Complexity and the Value of Thinking," Tech. rep., Institute of Labor Economics (IZA).

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

Kagel, J. H. and A. Penta (2021): "Unraveling in guessing games: An experimental study (by Rosemarie Nagel)," in *The Art of Experimental Economics*, Routledge, 109–118.

Kets, W. (2022): "Organizational Design: Culture and Incentives," *mimeo*.

Kets, W., W. Kager, and A. Sandroni (2022): "The value of a coordination game," *Journal of Economic Theory*, 201, 105419.

Kets, W. and A. Sandroni (2019): "A belief-based theory of homophily," *Games and Economic Behavior*, 115, 410–435.

——— (2021): "A theory of strategic uncertainty and cultural diversity," *The Review of Economic Studies*, 88, 287–333.

Lambrecht, M., E. Proto, A. Rustichini, and A. Sofianos (2022): "Intelligence Disclosure in Repeated Interactions," *working paper*.

Nagel, R. (1995): "Unraveling in guessing games: An experimental study," *The American Economic Review*, 85, 1313–1326.

Proto, E., A. Rustichini, and A. Sofianos (2019): "Intelligence, personality, and gains from cooperation in repeated interactions," *Journal of Political Economy*, 127, 1351–1390.

——— (2021): "Intelligence, errors and cooperation in repeated interactions," *The Review of Economic Studies*.

Raven, J. (1994): *Raven's Advanced Progressive Matrices & Mill Hill Vocabulary Scale*, Harcourt Assessment.

Schelling, T. (1960): *The Strategy of Conflict*, Harvard University Press, Cambridge, MA.

Sugden, R. (1995): "A theory of focal points," *The Economic Journal*, 105, 533–550.

Weber, R. A. (2001): "Behavior and learning in the "dirty faces" game," *Experimental Economics*, 4, 229–242.

A Appendix

A.1 Regression Tables for BoS Prediction 2

	(1) Choice of X	(2) Choice of X
High Payoff Dummy	0.151**	0.0189
	(2.21)	(0.38)
Constant	0.528***	0.349***
	(15.48)	(13.89)
Observations	212	212

t statistics in parentheses

* p¡0.10, ** p¡0.05, *** p¡0.01

Table 15: BoS Game: Panel regression results testing Prediction 2.1 (H label players) Model (1) gives the choice of preferred action X in the BoS game against a H label opponent while Model (2) gives the results against a L label opponent. Standard errors are clustered at the subject level.

	(1) Choice of X	(2) Choice of X
High Payoff Dummy	0.0698	-0.0233
	(0.68)	(-0.24)
Constant	0.535***	0.628***
	(10.43)	(12.88)
Observations	86	86

t statistics in parentheses

* p¡0.10, ** p¡0.05, *** p¡0.01

Table 16: BoS Game: Panel regression results testing Prediction 2.2 (L label players) Model (1) gives the choice of preferred action X in the BoS game against a L label opponent while Model (2) gives the results against a H label opponent. Standard errors are clustered at the subject level.

A.2 Regression Tables for RevSA Game

	(1)	(2)	(3)	(4)
	Choice of X	Choice of X	Choice of X	Choice of X
Opponent has H label	-0.160***	-0.160***	-0.0283	-0.0283
	(-2.66)	(-2.65)	(-0.54)	(-0.53)
High Confidence Test		-0.0251		0.0687
		(-0.30)		(0.81)
Constant	0.349***	0.370***	0.255***	0.198**
	(11.58)	(4.38)	(9.64)	(2.46)
Observations	212	212	212	212

t statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Table 17: RevSA Game: Panel regression results for H label players
Models (1) and (2) give the results for the low payoff versions of the RevSA game and (3) and (4) for the high payoff versions. Standard errors are clustered at the subject level.

	(1)	(2)	(3)	(4)
	Choice of X	Choice of X	Choice of X	Choice of X
Opponent has H label	0.0698	0.0698	-0.0465	-0.0465
	(0.57)	(0.57)	(-0.49)	(-0.49)
High Confidence Test		0.143		0.170
		(1.59)		(1.53)
Constant	0.395***	0.322***	0.372***	0.285***
	(6.45)	(3.84)	(7.88)	(3.18)
Observations	86	86	86	86

t statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Table 18: RevSA Game: Panel regression results for L label players
Models (1) and (2) give the results for the low payoff versions of the RevSA game and (3) and (4) for the high payoff versions. Standard errors are clustered at the subject level.

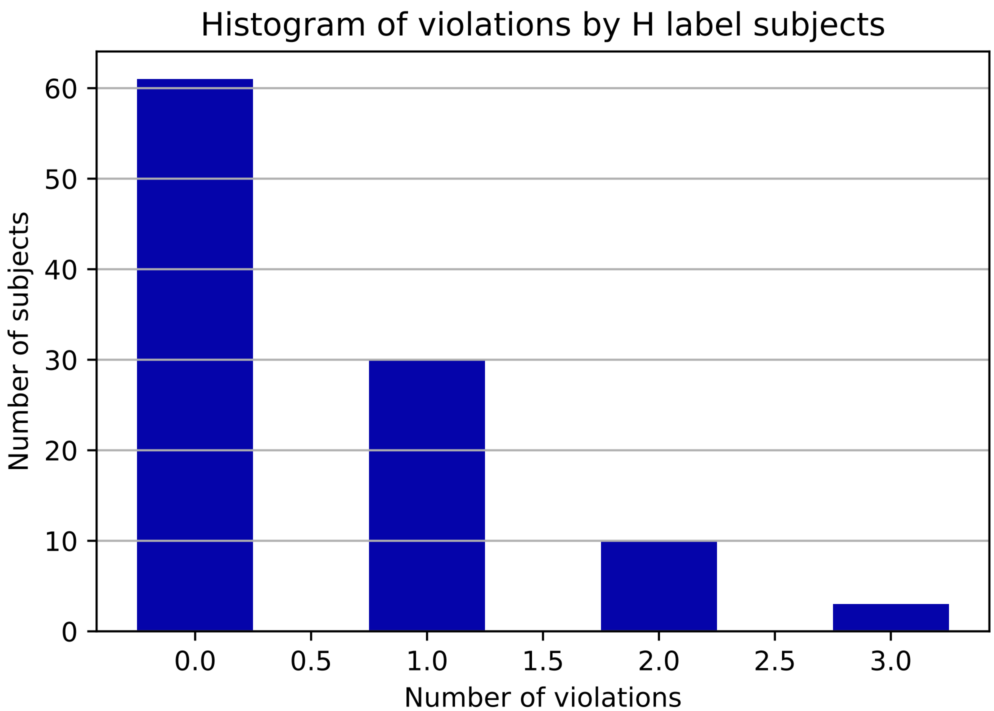

Figure 12: Histogram displaying the number of violations of H label subjects of conditions on behavior set out by the model

A.3 Additional Results

A.3.1 BoS: Violations of Theoretical Predictions

It is possible that the aggregate results that we have discussed for the BoS game are driven by a small number of subjects, or possibly result from individual effects different to those predicted by the model averaging out to look like the desirable outcomes at the aggregate level. To assess whether this is the case, we conduct an individual level analysis. For this analysis, we use the predictions of the model for the BoS game and assess whether subjects violate these. For the H label group, if a subject chooses X against a H label opponent in the low payoff version, they should not switch to playing Y when the payoff increases. Similarly, if a subject plays Y against a L label subject, thus expecting the equilibrium favorable for the L player to occur, they should not switch to playing X when payoffs increase. If a H subject already played Y against another H label player, they should keep doing so when the opponent changes to a L label player. If a subject violates any of these conditions, we count it as a violation and add the total number of violations for each subject. A histogram with the H label players' number of violations is given in Figure 12. The Figure shows that the number of violations is generally very low with nearly 90% of H label subjects having one or no violation. Individual level results are thus consistent with those at the aggregate level.

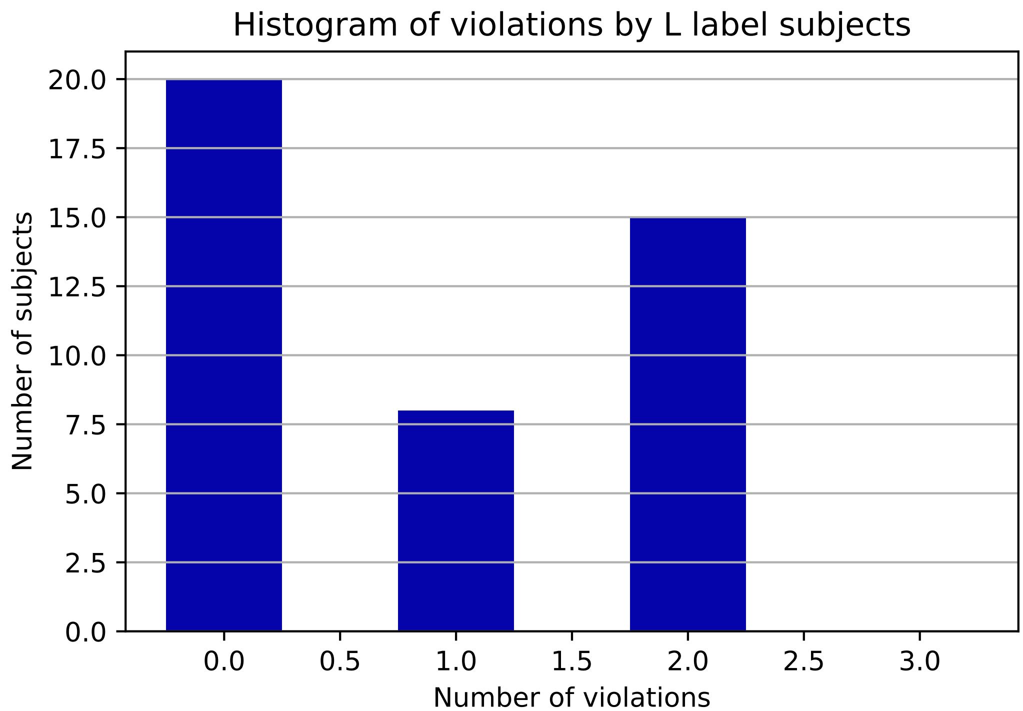

Figure 13: Histogram displaying the number of violations of L label subjects of conditions on behavior set out by the model

For the Low label subjects, if someone selected X against a L opponent for the low payoff version of the game, they should not switch to playing Y for any of the other versions of the BoS game. A histogram of the violations of the L label subjects is given in Figure 13. Here, around two thirds of subjects have one or fewer violations again suggesting that aggregate and individual results are comparable.

A.3.2 Raven test versus our cognitive test

In order to compare both tests, we first examine the correlation between the two tests. This is fairly high at 24.17%. Perhaps more interesting is whether resulting groups are correlated. Here, we find that the tests agreed on 24.18% of group allocations. This suggests that either the tests measure a different characteristic or that subjects were fatigued when they completed the second test at the end of the experiment, leading to inconsistent results.

Most importantly for our experiment, however, is the question of whether subjects believed in the tests' validity and thus in the labels of subjects. Here, we find that belief in the tests is very similar across both treatments (Our Test first versus Raven test first). A Kolmogorov-Smirnov test cannot reject the null that responses to the question about belief in the test come from the same distribution for both versions of the cognitive test (p-value = 0.834 for the Combined Kolmogorov-Smirnov test).

A.3.3 Unlabeled Treatment

Subjects who participated in the unlabeled treatments played the same BoS, RevSA, AMP and Stag Hunt games as the subjects in the labeled treatments. However, they were not informed of their own performance in the test or of that of their opponents. They completed two versions of each game, the low and the high payoff versions, against an unlabeled opponent (who was randomly drawn from the unlabeled group).

We find that 50% of subjects choose X in the low payoff version and roughly 47% in the high payoff version. This suggests that neither action is particularly salient.

For the RevSA game, 29.4% of unlabeled subjects play X in the low payoff game and close to 26.5% in the high payoff game. This suggests that subjects anticipate that their opponent is likely to choose W and best-respond by playing Y.

In the Stag Hunt game, nearly 80% of subjects choose X in the low payoff game and 82.35% in the high payoff version. This level is comparable to the behavior of subjects in the labeled treatments. Table 19 gives the percentages for all of the above games.

	BoS	RevSA	Stag Hunt
Low payoff	50.00%	29.41%	79.41
High payoff	47.06%	26.47%	82.35

Table 19: Results for BoS, RevSA and Stag Hunt - Unlabeled Players: % choosing X (their preferred outcome)

For the AMP game, we find that row players choose their preferred action X with close to 70% in the low payoff game. When the asymmetry increases, the frequency with which they choose X increases to roughly 76.5%. Column players have flat incentives and in the low payoff version, i.e. with the small asymmetry, they behave as if the game was symmetric in the sense that 50% pick either Z or W. In the strongly asymmetric version, however, more than 82% of subjects play Z, which is the best response if the opponent chooses X. As with the labeled treatments, this suggests that subjects react to changes in the value function of their opponents. Results are given in Table 20 below.

	Row	Column
Low payoff	70.59%	50.00%
High payoff	76.47%	82.35%

Table 20: Results for AMP - Unlabeled Players: % of row (column) players choosing X (Z)

A.4 Experimental Design

A.4.1 Experimental Structure

Before starting the experiment, subjects were randomly assigned the role of either row or column player. The subjects first completed either Our Cognitive Test or the Raven Test. Based on their performance, they were then assigned to the Low label, the High label or the Unlabeled group. Subjects first played the BoS game, then the Stag Hunt game, the RevSA game and finally the Asymmetric Matching Pennies (AMP) game. Due to their symmetric nature, players saw the games displayed as the row player's game; for the AMP game, the game was shown as either the row or the column version. Before each type of game, i.e. BoS, RevSA, Stag Hunt or AMP, subjects had to complete comprehension checks. The four versions of the games were played in the following sequence. First, the low payoff version against an opponent with the same label as them, then against an opponent with the opposite label. Second, the high payoff version against someone from the same and then from the other label group. After completion of the main games, subjects answered a question on how much they believed performance in the test was correlated with performance in the games. They then played the alternative test, i.e. either the Raven test or Our Cognitive Test, depending on which test they had already completed. Afterwards, they participated in a hypothetical 11-20 game, subjects in the labeled treatments played both against a hypothetical H and a L label player, and answered three questions from the Cognitive Reflection Test (Frederick (2005)).

A.4.2 Experimental Instructions

The experiment was conducted in Spanish as all participants were students at a Spanish university. The instructions displayed here are translations to English.

Instructions for Our Test

Note that the cognitive test contained the same questions as the cognitive tests used in Alaoui and Penta (2016) and Alaoui et al. (2020) and that instructions for the individual questions are thus identical. However, the test contained only three questions: the Mastermind game, the Centipede game and the Muddy Faces game. For the exact instructions for these games, see Alaoui et al. (2020). Similarly, the wording of the hypothetical 11-20 game is the same as used in Alaoui and Penta (2016) and Alaoui et al. (2020) with the only difference being that it is hypothetical.

Scoring of Our Cognitive Test

The Mastermind game gave a total of 100 points if the correct sequence was entered. Otherwise, subjects received 15 points for a correct number in the correct place and 5 points for a correct number in the wrong place, in the last round. The Centipede game gave a total of 100 points if the correct answer was given. Otherwise, they received 60, 45, 30, 15, or 0 points depending on how close their answer was to the true one. For the Muddy Faces game, subjects obtained 120 points if each sub-question was correctly answered. Alternatively, they received partial points depending on how closely their reasoning followed the correct iterative reasoning. The points were summed up and divided by 3.2 to create a maximum of 100 points.

Instructions for Raven Test

Subjects completed Set I to keep the time of the experimental sessions below two hours. They completed a practice question to show what it means for a piece to be "correct" in the sense that it completes the pattern shown on the screen. Instructions for each of the twelve questions were the following:

Please select the correct piece from the eight pieces shown below. You can select the piece by clicking on the corresponding number.

Instructions for BoS, RevSA, Stag Hunt, and AMP games

Your score in the test was: very high (low).
The other player is:
- A person with a very high (low) score in the test.
- Furthermore, they have the same information as you.

To make your choice, click on one of the buttons.

The game matrix was displayed below the text. The specific matrices of each game can be found in the main text. Each version of a game was shown on a separate screen.

Instructions for "Belief in Test" question

Please indicate to what degree you agree with the following statement, on a scale from 1 (I do not agree) to 5 (I fully agree):

"A higher score in the test indicates that the person can more easily reason in the games of this experiment."

Instructions for Hypothetical 11-20 game

Imagine a game with following structure:

Pick a number between 11 and 20. You will always receive the amount that you announce, in tokens.

In addition:

- If you give the same number as your opponent, you receive an extra 10 tokens.
- If you give a number that's exactly one less than your opponent, you receive an extra 20 tokens.

Imagine that your opponent is someone who:
- has a low/very high score in the test.
- has been given the same rules as you.

Instructions for CRT questions

For the wording of the three CRT questions, please see Frederick (2005).