Coordination and Sophistication

Larbi Alaoui, Katharina A. Janezic, Antonio Penta

April 2022

Abstract

How coordination can be achieved in isolated, one-shot interactions is a long-standing question in game theory. Without communication and in the absence of focal points, whether coordination can be reached at all is unclear. We show, however, that in a non-equilibrium model in which the level of reasoning responds to incentives, high stakes may induce coordination when the cognitive sophistication of the players is heterogeneous and when this is agreed upon. The equilibrium on which coordination is expected to occur, according to our model, depends on the payoff structure of the game in ways that differ from those implicit in standard solution concepts, or from the implications that one could draw applying exogenous criteria for the attribution of the strategic advantage. Our model therefore provides a novel mechanism for endogenous coordination and one in which it is differences between players, rather than their similarities, that lead to increased coordination. Using the model as a framework, we conduct an experiment to examine coordination in such a setting.

1 Introduction

The vast majority of games analyzed in economics contain some aspect of coordination, and yet the question of understanding how coordination can be achieved in isolated settings remains open. One proposed mechanism uses a notion of focal points (Schelling (1960)), which depends first on shared culture, since there must be a common view concerning which points are focal, and second on there being focality within the specific environment of interest. In practice, and especially when agents face novel strategic situations, these conditions are often not met. In such settings it is unclear how, or even whether, coordination is achieved. In this paper we propose a novel mechanism for how coordination can be reached, in the absence of focal points, in environments that inherently feature strategic uncertainty. We then conduct an experiment to test our predictions, and find that observed behavior is indeed consistent with the mechanism proposed.

To fix ideas, consider the canonical battle of the sexes (BoS) game, in which two players both prefer coordinating on the same location (action profile), but have different preferences over which place to coordinate on. If neither location is focal, if there is no repetition, and no communication, then it could appear difficult to see how these agents can coordinate. Now suppose that the players have beliefs over each other's 'cognitive sophistication' (or 'reasoning ability'), and consider the following two situations. In case (i), the agents view each other to be of similar sophistication, and in case (ii) they believe that one player is of higher cognitive sophistication than the other. Such situations take place frequently, and it seems intuitive that beliefs over relative sophistication may matter in behavior (see, e.g., Proto, Rustichini, and Sofianos (2019, 2021) for evidence on this). But little is known about the mechanism behind why such beliefs should matter for coordination and what behavior to expect, especially in the context of one-shot, isolated interactions, when there are no clear focal points.

In the mechanism we propose, these beliefs over cognitive sophistication have clear implications concerning the occurrence and the nature of coordination. Perhaps surprisingly, and in contrast with related approaches to coordination in different settings (Kets and Sandroni (2019, 2021) and Kets, Kager, and Sandroni (2022)), in the environments that we consider our model predicts that it is *heterogeneity* between players, rather than *homogeneity*, that favors coordination. In particular, in the situations described above, our model predicts higher coordination in case (ii), when players believe that they have different sophistication, than in case (i), when they are of similar sophistication. Furthermore, under this mechanism, players in the BoS game are more likely to coordinate on the equilibrium that is most favorable to the player who is believed to be *less* sophisticated. In other words, and in contrast to what other theories or intuition may suggest for this game, according to the mechanism we propose being perceived to be more sophisticated is a disadvantage when players face strategic uncertainty in the BoS game. Finally, and in contrast with what one might expect taking into account standard models of social preferences, according to our model, payoff transformations that exacerbate the disagreement between players in the BoS game (while preserving the symmetry of the game), do not reduce coordination, and may in fact favor its occurrence.

In essence, the logic of our model is the following. As each player tries to understand the situation and the other player's reasoning, his reasoning carries him or her deeper into considering what the other player might do. The stopping rule that we adopt for the players' reasoning process is based on the axiomatic foundation for general cost-benefit criteria in reasoning of Alaoui and Penta (2022). In the case of the BoS game, these criteria imply that if a player's favorite equilibrium is sufficiently preferred over the other one, then his reasoning is more likely to eventually end there (this is not true for other games of coordination, as we explain below, or for reasoning processes that 'stabilize', as those considered in Kets and Sandroni (2019, 2021) or those that arise in the presence of focal points). If the two players perceive each other to be of comparable sophistication, or if they do not agree on their relative sophistication, they would both play according to their maximum understanding, and hence typically end up in a miscoordination outcome. But when there is agreement that one player is sufficiently more sophisticated than the other, then the one who is perceived to be more sophisticated (higher type) believes that he has gone deeper in the chain of reasoning than the lower type. This added understanding of the situation then leads the higher type to play according to his belief over the lower type's reasoning. This induces him to 'yield' and attempt to coordinate on his less preferred equilibrium. As for the lower type, he still follows his own understanding and attempts to coordinate on his own most preferred equilibrium. Hence the result.

We stress a few important features of this logic: First, this mechanism depends on players' beliefs about their *relative* sophistication alone, regardless of their actual sophistication (that is, if players *commonly agree* that player 1 is more sophisticated than 2, then they would coordinate on 2's favorite equilibrium in the BoS game, even if 2 is actually more sophisticated than 1). Second, we will demonstrate that this logic holds for various potential forms of reasoning, and is not tied to a specific kind of chain, be it level-k reasoning or selection between possible equilibrium profiles. We will formally provide the crucial assumption required, which is a notion of *responsiveness* of the path of reasoning. Third, the players' belief over the others' form of reasoning itself need not be correct either. The proposed mechanism is therefore robust in the sense that it does not rely on players to be exactly correct about each other.

It may seem that it is the asymmetry in player labels itself that helps achieve more coordination. That is, if we were to label the higher sophistication agents "High" and the other "Low", then one view is that they would use "High" as a coordination device and play according to that type's preferred action, effectively replacing a theory of focal points with one of focal labels. But this is not what our model predicts. According to our explanation, coordination in the BoS class of games is more likely to occur on the preferred profile of the *low* types, not the high. Moreover, as we discuss below, our model has additional, more subtle implications that would arguably not be expected with a theory of 'focal labels'. For instance, in addition to the predictions discussed previously, our model also predicts an asymmetry between the high and the low types in that the high type are more likely to coordinate on the preferred action profile of the opponent when playing against low than high, while the low types are equally likely to play according to their own preferred profile in both 'treatments'. Finally, our model predicts that the high type concedes more to the low type in some classes of coordination games but expects the low type to concede more in others, which is difficult to reconcile with the idea that the channel is label focality.

The argument above may appear to suggest 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 capture adequately the logic of our model. To illustrate the difference, we propose another coordination game, which we refer to as the *modified BoS*, in which our model delivers the opposite prediction to the one obtained from the 'first mover' argument (Figure 1 shows examples of the two games side by side). In particular, in the modified BoS, 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 induces players' reasoning *not* to stop at the low types' favorite equilibrium, but rather at the high type's. For this reason, unlike in the case of 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. 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 type (these points will be discussed in detail below).

BoS Game

	W	Z
X	$70, 50$	$0, 0$
Y	$0, 0$	$50, 70$

Modified BoS Game

	W	Z
X	$130, 130$	$230, 220$
Y	$220, 230$	$170, 170$

Figure 1: Examples of the BoS and Modified BoS Games

After introducing our model and formally providing our theoretical results, we conduct a laboratory experiment to test the predictions. After taking a test of strategic sophistication, the subjects are labeled according to their scores. The higher and lowest scoring subjects play both against their own labels and against the other in the BoS game, under different incentive structures, which we refer to as 'low payoff' and 'high payoff'. Subjects also play the modified BoS discussed above. Our main identification assumption is that (i) subjects commonly agree that those who receive a high label have higher sophistication than those who have a low label, and (ii) subjects who are labeled in the same way commonly agree that they have similar cognitive sophistication as each other.[1]

The empirical findings are in line with the predictions of the model. Specifically, in the *canonical BoS*, we find that: (i) the high types concede more against low than against high; (ii) this effect is more pronounced for the high payoff treatments than low payoffs treatments; (iii) low types concede in a similar manner against low than against high; (iv) there is more coordination when playing against other types than against own type; (v) the increased coordination occurs on the low type's favorite equilibrium; and (vi) this effect is stronger for the high payoff treatments. Put together, this set of results lends support to the empirical relevance of the mechanism introduced by our model.

We then use the *modified BoS* to disentangle the predictions of our model from those of the alternative mechanisms discussed above, and in order to gain further understanding about the reasoning of our subjects. We find that the results are not in line with the view that low types obtain a first-mover advantage, and so it does not appear that this could be a main alternative explanation for our results above. As for the notion of 'label focality' that we mentioned above, we note that its implication is instead inconsistent with the results for the BoS game, which go in the opposite direction. We then analyze what these additional results indicate from the perspective of our framework, and show that they are overall consistent with the mechanism discussed above, under some further intuitive assumptions concerning the subjects' underlying path of reasoning. This exercise serves to illustrate how our approach can be used to gain some insight into subjects' reasoning process.

Lastly, we consider additional games as robustness checks of our results. In particular, we consider a stag hunt game and an asymmetric matching pennies game. We again find consistency between the results and our predictions, for both games. We further find that our model predicts behavior that risk-dominance does not, lending support to the mechanism introduced here.

In brief, we introduce a model of strategic reasoning that addresses one of the central issues of game theory, namely that of how coordination can be achieved in isolated interactions and in the absence of focal points. We find that subjects' behavior in a laboratory experiment is consistent with the predictions of this model. Abstracting from the framework suggested in this paper, the experimental results are in themselves novel. We also consider whether these findings could be due to merely providing a coordination device through labels, and show that this is not the case. We consider as well whether our model can be viewed as effectively granting the low type a first-mover advantage by reasoning less, and show that this is not the case theoretically, nor is it consistent with our further empirical findings. Moreover, we show that, perhaps contrary to a common perception, it is not necessarily the case that being of higher cognitive sophistication is beneficial to the agent, in the face of strategic uncertainty: while it may often be the case, in settings such

[1] If this assumption were not to hold and the subjects believe the test is uninformative, then according to our model, we would find no change in behavior. After they play the games, we ask subjects whether they find the test to be informative of strategic sophistication, and find that a high fraction does.

as the BoS game considered, the opposite is true.

The rest of this paper is structured as follows. Section 2 presents the model and theoretical predictions; Section 3 discusses the experimental design; Section 4 analyzes the main experimental results of the experiment; Section 5 discusses the additional results and Section 6 concludes.

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 $u_i : A_1 \times A_2 \to \mathbb{R}$ 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 parametrized by $x \in \mathbb{R}$, $x \geq 1$:

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

Table 1: 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 labels of the actions are focal or salient in some way, then focal point theory does not provide guidance as to how coordination can be achieved, if at all. Our focus here will be on this case, which also conveys the main logic for the model's implications for the other games in our experiments.

We assume that both players follow a stepwise process in their deliberation in the game, and that their understanding of the game – namely, the number of steps of reasoning they perform – is driven by a cost-benefit analysis, that 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, the players form beliefs about the opponent's understanding, and choose the action that maximizes their expected utility, given their beliefs about the opponent's that have resulted from their reasoning. We introduce next the model of such a deliberation process, introducing in order its main elements: the path of reasoning; the stopping rule; the 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, $a_j^{i,k} \in A_i$ represents i's best conjecture at step k about how an opponent he regards as at least equally sophisticated as he is, and let $a_i^{i,k} \in BR_i(a_j^{i,k})$, denote his best response to that conjecture, where we let $BR_i : A_j \rightrightarrows A_i$ denote player i's pure-action best reply correspondence, $BR_i(a_j) := \arg\max_{a_i \in A_i} u(a_i, a_j)$ for all $a_j \in A_j$.[2]

The path of reasoning $\{(a_i^{i,k}, a_j^{i,k})\}_{k\in\mathbb{N}}$ represents the sequence of conjectures and behaviors 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. Rather, we view reasoning as costly, and at any given step k players may well decide that it is not worth continuing reasoning. 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 will 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 (DOE): 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 would correspond 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 a current step, a player believes that an equilibrium is the one that an opponent's reasoning would lead him to choose. It may be that as he thinks more, he remains convinced of this equilibrium. In that case, his path of reasoning would 'stabilize' over that equilibrium profile. But 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 action at step k be equal to the action of an opponent of level (k-1). Formally, for each $k = 1, 2, ...,$ $a_j^{i,k} = a_j^i(k-1)$, where for an arbitrary level-0 *anchor* $a^i(0) = \left(a_1^i(0), a_2^i(0)\right)$, we have $a_l^i(k) = BR\left(a_{-l}^i(k-1)\right)$ for each $l \in \{1, 2\}$. Note

[2]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.

that, for this specific model, if the anchor $a^i(0)$ is a Nash equilibrium $a^* \in A$, then the path of reasoning is *constant*, in the sense that $a^i(k) = a^*$ for all k: in this case, in his deliberation player i only contemplates playing the action $a_i^{i,k} = a_i^*$ at any step k. Within the level-k mode of reasoning, the situation of a Nash equilibrium anchor can thus be thought of as a situation in which player i subjectively believes that there is a social norm which prescribes to play a^*, and further reasoning about mutual best replies does not challenge such initial disposition. If, in contrast, $a(0)$ 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 $a_i^i(k)$ 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 of the sticky, *absorbing* type, 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 DOE mode of reasoning, this would be if the reasoning sequence does not *ever* stabilize on any one equilibrium; in the case of level-k reasoning, this would be 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 $k \geq \bar{k}$, $a_i^{i,k} = a_i^{i,k+1}$. 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 the $\bar{k}$ after which it no longer changes. In the case where such $\bar{k}_i = 0$, then 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's (2019, 2021) *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, 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 of responsiveness 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's (2019, 2021) papers, since there k is taken to infinity), they would potentially always question 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 i 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 benefits instead only depend on the game's payoffs, such as the x parameter in the BoS game. When deciding whether or not to reason at that step, the agent compares the two, and continues so long as the value of reasoning exceeds the cost of reasoning, i.e., so long as $v_i(k) \geq c_i(k)$. For later 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. Player i's *cognitive bound* is the value that this function takes at (c_i, v_i):

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

We say that cost function c' is 'more (resp. less) sophisticated' than c'', if $c'(k) \leq c''(k)$ (resp., c'', if $c'(k) \geq c''(k)$) for every k. For any $c_i \in \mathbb{R}_+^{\mathbb{N}}$, we denote by $C^+(c_i)$ and $C^-(c_i)$ 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 to 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 *maximum gain* 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 qualitative similar implications to the ones we discuss in the following. Hence, while it can be relaxed without significantly affecting the results, the representation has the advantage of having no free parameter and hence it offers no degrees of freedom. This representation of the cost of reasoning will therefore be maintained throughout.

2.3 Beliefs about Others' Reasoning and Choice

The depth of reasoning describes the thought process of the agent, but his behavior also depends on his beliefs over his opponent, and particularly over the opponent's cost function. Such beliefs are then used to derive i's beliefs about 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.[3] Player

[3] 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

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^i_j, and taking into account j's value of reasoning, as entailed by i's own understanding of j's reasoning. Formally, let $v^i_j : \mathbb{N} \to \mathbb{R}$ be such that

$$v^i_j(k) = \max_{a_i \in A_i} u_j(BR_j(a_i), a_i) - u_j(a^{i,k-1}_j, a_i).$$

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

$$\hat{k}^i_j = \min\left\{\hat{k}_i, \mathcal{K}(c^i_j, v^i_j)\right\}. \tag{3}$$

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

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 depth of reasoning.

The simple model above of course is not meant to capture the precise behavior that would be observed empirically, but rather to illustrate the mechanisms of the model, which we will show has very sharp predictions in the BoS game, which overall will provide a force towards increased coordination on the preferred action profile of the less sophisticated opponent when there is heterogeneity of costs. In practice, it is reasonable to consider various extensions.

2.4 Endogenous Deliberation in the BoS game

Consider player 1's value of reasoning in the baseline BoS game of Table 1. When $a^{1,k-1}_1 = B_1$, then $v_1(k) = \max\{x - x, 1 - 0\} = 1$, and when $a^{1,k-1}_1 = W_1$, then $v_1(k) = \max\{x - x, x - 0\} = x$. There is thus an asymmetry between the two: if, at step $k-1$, the player believes that B_1 is best, then the maximum gain he could obtain is only 1; But if he believes that W_1 is best, then he has more to gain, and his value is now x. If x increases, the maximum gain at step in which $a^{k-1}_1 = B_1$ is not affected and remains at 1, while it increases at steps in $a^{1,k-1}_1 = W_1$. 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 $a^{1,k-1}_1 = a^{1,k}_1 = B_1$ or, alternatively $a^{1,k-1}_1 = B_1$ and $a^{1,k}_1 = W_1$, the value of reasoning for the k-th step is then the same for both. This is because the agent

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)). We will discuss below how the results that we obtain based on the *first-order types* introduced in the main text would extend to more general hierarchies of beliefs over cost functions.

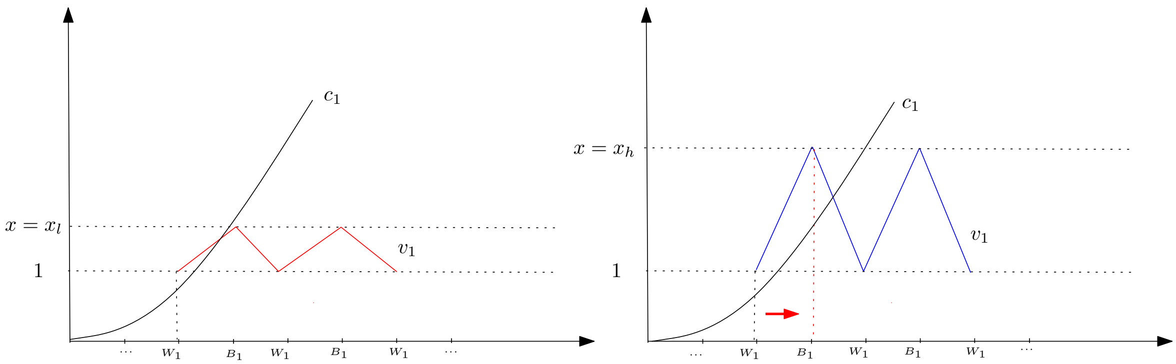

Figure 2: Low-payoff depth of reasoning such that $a_1^{1,\hat{k}_1} = W_1$.

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, then for any x in the BoS game, for any player i and for his associated path of reasoning, there is a $\hat{k}_i(x)$ for which $c_i(\hat{k}_i) > v(\hat{k}_i)$, which is the stopping rule for player i at that x. This simple structure yields very sharp implications for any path of reasoning that is *responsive*.[4] For any such path of reasoning, and for any x, consider any player i with a responsive path, and whose last step of reasoning is $\hat{k}_i(x)$. 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 conduct the $(\hat{k}_i + 1)$-th step, $c_i(\hat{k}_i + 1) > 1$. An increase in x has no effect on $v_i(\hat{k}_i + 1)$, and so the threshold $\hat{k}_i(x)$ remains unchanged as x goes up. Now suppose instead that $a_i^{i,\hat{k}_i} = W_i$. Then $v_i(\hat{k}_i + 1) = x$, and $c_i(\hat{k}_i + 1) > x$. Since $c_i(\hat{k}_i + 1)$ is not infinite, then for high enough x', $x' > c_i(\hat{k}_i + 1)$, and the agent would then 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 x', 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 $c_i(\tilde{k}_i + 1) > c_i(\hat{k}_i + 1) > x > 1$. 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 $\bar{x}$ such that, for all $x > \bar{x}$, player i stops reasoning at some step $\hat{k}(x)$ such that $a_i^{i,\hat{k}_i(x)} = B_i$*

The logic of this result is illustrated in Figures 2 and 3, in which a responsive path is taken to be one for which $a_i^{i,k}$ 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

[4] 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.

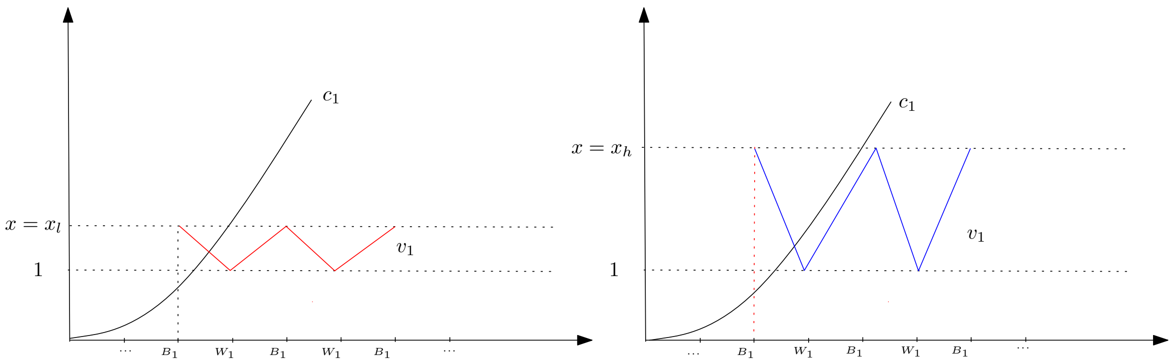

Figure 3: Low-payoff depth of reasoning such that $a_1^{1,\hat{k}_1} = B_1$.

either (B_1, B_2) or (W_1, W_2), and where the path then alternates between (B_1, B_2) and (W_1, W_2). It would also be the case for equilibrium selection reasoning, if the player alternates between (B_1, W_1) and (B_2, W_2) as the equilibria.

As can be seen in Figure 2, if 1's depth $\hat{k}_i$ for lower x (taken to be x_l) has associated $a_1^{1,\hat{k}_i} = W_1$, then a high enough increase in x (from x_l to x_h, in the figures) will lead to B_1. If, as in Figure 3, 1's depth $\hat{k}_1$ for lower x has associated $a_1^{1,\hat{k}_1} = B_1$, then an increase in x has no effect. Whereas the actual step $\hat{k}_1$ 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 x.

Note that applying the logic above 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 $\bar{x}$ such that, for all $x > \bar{x}$, player i thinks that j stops reasoning at some step $\hat{k}_j^i(x)$ such that $a_j^{i,\hat{k}_j^i(x)} = B_j$.*

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, however, the inequality is weak, because it may be that the cost functions are very close to each other, and hence for some incentives they would effectively entail the same depth of reasoning. The next assumption rules out this possibility, in that it requires that players' beliefs about the opponent's sophistication is effectively different than one's own, in the sense that beliefs are sufficiently lower (resp., higher) to effectively entail different depths of reasoning.

Formally: Fix player i's path of reasoning in the BoS game, and type $t_i = (t_i, c_j^i)$. 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 $x \geq 1$, $\mathcal{K}(v_i, c_i) < \mathcal{K}(v_i^j, c_i^j)$ (resp., $c_j^i \in C^-(c_i)$ and $\mathcal{K}(v_i, c_i) > \mathcal{K}(v_i^j, c_i^j)$).

Lemma 3 *Under the maintained assumptions, for any* responsive *path of reasoning, in* *the BoS game above there exists $\bar{x}$ such that, for all $x > \bar{x}$, 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.*

Proposition 1 *If both players agree that i is strictly more sophisticated than j, and if their paths of reasoning are* responsive*, there exists $\bar{x}$ such that, for all $x > \bar{x}$, players play $a = (W_i, B_j)$, the Nash equilibrium most favorable to player j.*

2.5 Focality, Alignment and Eductive Coordination: Discussion

Since Schelling (1960), a *focal point* is an outcome that is salient (i.e., agents expect it to happen) and self-enforcing (so that further inspection confirms the initial expectation). Hence, if such a focal point exists – be it due to payoff considerations (e.g., if efficiency, risk-dominance, etc. are shared refinement criteria) or 'non mathematical' properties of the game (e.g., intrisic characteristis or labeling of actions, cf. Crawford et al. (2008), or Sugden (1995),

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.*

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 does not believe in a 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 playing the obvious coordination game against each other. If not aware of the nationality of the opponent, they would (most likely) 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. In this thought experiment, it is natural to hypothesize that if the two drivers were each made aware of the nationality of the opponent (and this was CK), then the individual $a^{i,1}$ 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 be an instance of the first case.

The results below will show 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: (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 x is sufficiently high); (ii) second, if both players agree on their relative sophistication – that is: it is, if they commonly believe (albeit wrongly so) that player i is more sophisticated than player j.

3 Experimental design and predictions

3.1 Experimental design

The experiment was designed to test whether coordination increases for some games, such as the BoS, when players of different sophistication play against each other rather than players of the same sophistication. We included two main games that allow to test the predictions of the model as well as two additional games to test alternative mechanisms. The predictions require that subjects indeed have different levels of sophistication and believe that these levels exist. Subjects 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 beginning of the experiment, we therefore let all subjects complete a test of cognitive sophistication. The test contained the Muddy Faces game (see for example Weber (2001)), a version of the Mastermind game and a centipede game. The questions were the same as in Alaoui, Janezic, and Penta (2020) and in Alaoui and Penta (2016).

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 (APM test) saw the APM test (resp. our test) at the end).

We then separate the subjects into three groups, which we denote as High, Moderate, and Low. For the main experiment, we use only the High and Low groups, so as to obtain enough perceived distance between the groups. The High and Low groups are informed of their labels. The Moderate group plays an unlabeled treatment and is not informed of their labels or of those of their opponents. The cutoff expectations were based on the distributions of Alaoui and Penta (2016), Alaoui, Janezic, and Penta (2020) and subsequent pilots.[5]

The first game that subjects played is the BoS game where we predict increased coordination, on the preferred action profile of the less sophisticated player, when players of different sophistication play together compared to players of the same sophistication. The BoS game was of the following kind, both against their own label and against the label of the opponent:

[5]Cutoffs are not determined session by session. Subjects of the entire sample played against one another, and were paid once the sessions ended.

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

Table 2: Battle of the Sexes Game

where $r \in \{51, 70\}$, depending on the treatment. The labels, X and Y, W and Z were chosen to be relatively neutral in their salience. Below, we will use the language B_i and W_i rather than X (Z) and Y (W), respectively, for ease of mapping with the predictions above. Each subject played the following four versions of the game, without feedback and with random, anonymous matching at every round:

- **BoS-Own-S:** The BoS game is played against someone from one's own label, for the smaller $r = 51$.
- **BoS-Other-S:** The BoS game against someone from the other label for $r = 51$.
- **BoS-Own-G:** The BoS game against someone from one's own label for the greater $r = 70$.
- **BoS-Other-G:** The BoS game against someone from the other label, for $r = 70$.

In addition to the BoS game, subjects also played a game, denoted as Modified BoS, specifically designed to test whether the findings from the BoS game can be attributed to first-mover advantage rather than to our model. The game takes the following form:

	W	Z
X	$130, 130$	$230, r$
Y	$r, 230$	$170, 170$

Table 3: Modified BoS Game

where $r \in \{190, 220\}$, depending on the treatment. As with the BoS, the subjects played four versions of this game: with $r = 190$ against someone from their own label or against someone from the other label and with $r = 220$ against someone from their label or from the other label.

The two additional games which assess the viability of some alternative mechanism are explained in detail in Section 5.

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

3.2 Predictions for the Experiment

The propositions from Section 2 map to testable predictions for the experiment, under the following assumptions. Suppose that those labeled Low score in the experiment are of type c_H (low sophistication) and that those with a high score are of type c_H. Moreover, suppose that a non-zero fraction of subjects of each type have a responsive path, and that for those subjects, $r \in \{51, 70\}$ are over the $\bar{r}_i$ threshold for those paths, for $i\ in\{1, 2\}$.[6]

Let $p^a(\dot{})$, for $a \in \{L, H\}$ (L being the low group, H being the High group) be the percentage of subjects of type a playing B_1 or B_2) (i.e., X or Z, in the BoS game given to the subjects), where the argument in the function refers to the treatment.

Prediction 1 *The following predictions for the BoS game follow from our model:*

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

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

The model also gives the following predictions for the Modified BoS game.

Prediction 2 *The predictions for the second game are as follows:*

1. $p^H(\text{Modified BoS-Own-G}) \leq p^H(\text{Modified BoS-Other-G})$*: the percentage of High subjects playing their own preferred action in the Modified BoS game with sufficiently high values of* r *is* higher *when playing the other label than their own.*

2. $p^L(\text{Modified BoS-Own-G}) = p^L(\text{Modified BoS-Other-G})$*: the percentage of low subjects playing their own preferred action in the Modified BoS game with sufficiently high values of* r *is* the same *when playing the other label than their own.*

3.3 Logistics

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, 181 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 specification of each game i.e. one out of the four specifications 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. 147 subjects participated in the labeled treatments, of which 43 were classified as Low score type and 104 as High score type; 34 subjects were in the unlabeled group.

[6] The BoS game in Section 2 was described with different payoffs; the results for $\bar{x}_i$ trivially carry over to analogous results for $\bar{r}_i$.

4 Results

4.1 Test of Main Hypothesis - BoS Game

In this section, we discuss the results relating to Prediction 1, which is tested using the BoS game. We will first discuss the hypotheses relating to the belief effects and will then move on to the payoff effects. In the following, the player type of interest is assumed to be of the row player type and their opponents of the column player type.

The model predicts that High label players adjust their behaviour in the BoS game when playing against Low label players such that they play their own preferred action relatively less. In the BoS game, the preferred action of the row players is to choose X. For the High label players, in 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. However, when they face a Low label player, this percentage drops to 36%. 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 whether they are playing against a High or Low label opponent is significant at the 1% level (p-value= 0.008). We also compare the distribution of chosen actions using a Wilcoxon signed rank test. The p-value of the test statistic is 0.0004. When we repeat the analysis for the high payoff version of the game, we find that 69% play X against a High opponent but that only 37% play X against a Low opponent. The regression coefficient is significant at more than 0.1% (p-value$<$ 0.001). 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, less when they play against a Low label opponent than against a High label opponent. This supports Prediction 1.1.

Under the assumption of sufficiently high payoffs, the model predicts that the overall level of coordination should increase for heterogeneous relative to homogeneous labels. Table 4 shows the percentage with which each of the four cells of the game occurred for the High label players, for each of the four versions of the BoS game. For those versions of the game where High label players played against High label players, the sample was split along the row and column players. For those versions of the game where the High label players were matched with Low label players, the High label players take the role of the row player. Comparing Table 4(a) with (b) shows that coordination increases, from 50.4% to 53.69%, when the labels differ between players. For the high payoff version of the BoS game, this increase is larger, from 42.67% to 52.82%. This increase in the overall level of coordination when labels change from being homogeneous to heterogeneous supports the predictions of the model.

For the Low label group, in the low payoff BoS game, we find that 54% of subjects play their preferred action, X, against a Low label opponent. This percentage increases to

Table 4: Results BoS Game - High Label Players

(a) Low Payoff - Opponent has H label

	W	Z
X	30.95%	28.66%
Y	20.97%	19.45%

(b) Low Payoff - Opponent has L label

	W	Z
X	13.24%	22.34%
Y	23.97%	40.45%

(c) High Payoff - Opponent has H label

	W	Z
X	19.41%	47.89%
Y	9.43%	23.26%

(d) High Payoff - Opponent has L label

	W	Z
X	14.45%	22.09%
Y	25.09%	38.37%

63% when playing against an opponent from the High label group. This difference is not statistically significant (p-value= 0.421). The p-value of the Wilcoxon signed rank test statistic is 0.414. This result is consistent with the assumption that Low label players view other players as having a cost which is equal to their own. The small, but not significant, change in the percentage of the chosen preferred actions is consistent with a low number of Low label players viewing other Low label players as having a higher cost than themselves. For the high payoff version of the game, we find that 61% of Low label players choose their preferred action, irrespective of the label of their opponents. The regression coefficient is not significant and the Wilcoxon signed rank test statistic is also not significant (both have a p-value= 1). This is again consistent with Low label players believing that all other players are at least as sophisticated as themselves, supporting Prediction 1.2.

Table 5 shows the outcomes of each version of the BoS game for the Low label group. The percentages with which each cell occurred are calculated analogously to the High label group (see above). We find that, as with the High label group, the level of coordination increases when labels change from being homogeneous to being heterogeneous. For the low payoff version of the game, coordination increases from 50% to 53.69%. For the high payoff version of the game, the level of coordination increases from 50% to 52.82%. As with the High group, this increase in the level of coordination supports the model's predictions.

Table 5: Results BoS Game - Low Label Players

(a)Low Payoff - Opponent has L label

	W	Z
X	28.00%	28.00%
Y	22.00%	22.00%

(b) Low Payoff - Opponent has H label

	W	Z
X	40.45%	22.34%
Y	23.97%	13.24%

(c) High Payoff - Opponent has L label

	W	Z
X	34.00%	34.00%
Y	16.00%	16.00%

(d) High Payoff - Opponent has H label

	W	Z
X	38.37%	22.09%
Y	25.09%	14.45%

Under the assumption that the cognitive bound is binding for at least some subjects, we expect to find that the size of the predicted change in behaviour weakly increases when payoffs increase. For the High label players, this implies that we should see a larger difference in the frequency with which X is played across opponent labels in the high payoff version of the BoS game than in the low payoff version. Indeed we find that the difference between the frequencies increases from 18 percentage points in the low payoff version to 32 percentage points in the high payoff version. For the Low label players, the model does not predict a change in behaviour across labels such that we also do not expect a difference in the size of the effects. We find that for both the low and the high payoff versions the difference in behaviour is not significant, which is fully in line with the model's predictions.

Finally, note that the results for the BoS game are consistent with the presence of a noise type.

4.2 Additional Tests - Modified BoS Game

It is possible that the finding that subjects coordinate on the equilibrium most preferred by the Low label player is due to the Low label conferring a kind of first-mover advantage, in the sense that the Low label players commit to stopping first and High label players best respond to this. To test whether this is the case, we let subjects play the Modified BoS game, for which the first-mover advantage theory predicts that subjects act such as to coordinate on the most preferred action of the Low label player as well.

For the High label group, 19% of players choose their preferred action, X, when playing against someone with the High label in the low payoff version of the Modified BoS game. When the opponent changes to being a Low label player, this increases to 36%. We again conduct panel regressions 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 (p-value= 0.009). As with the BoS game, we also conduct a Wilcoxon signed rank test to confirm whether the two distributions of actions is statistically significant when the opponent's label changes. We obtain a p-value of 0.003. 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 high payoff version of the game, we find that 23% choose their preferred action when playing against another High label player. This percentage increases to 25% when the opponent changes to being a Low label player. However, this increase is not statistically significant (p-value of 0.7 in the regression and of 0.49 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, here.

For the Low label players, 40% choose their preferred action when playing against another Low label player while 47% select their preferred action against a High label opponent. This difference is not statistically significant (p-value= 0.57 in the regression and p-value= 0.3 for the Wilcoxon test). In the high payoff version, 37% select their preferred action against a Low label opponent and 33% against a High label opponent. Again, this difference is not statistically significant (p-value= 0.623 for the regression and p-value= 0.371 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.

While we can use the Modified BoS 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, such that it cannot make a prediction for the low payoff version of the Modified BoS 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, behaviour changes in the desired direction but that the change is small. Observed behaviour is also consistent with the existence of beliefs over noise players. Assuming that there exists a noise type with probability p, who plays either W or Z with 50 % probability, a row player should increase their likelihood of playing Y. 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 Results

5.1 Additional games: Stag Hunt

In the Stag Hunt game, we find that a large majority of subjects choose X, which is their preferred action. For the High label players in the low payoff version of the game, we find that 81% choose X when facing a High label opponent while 75% select X against a Low label opponent. For the high payoff version, we find similar results in that 86% select X against a High label opponent and 81% against a Low label opponent.

For the Low label players in the low payoff version, we observe that 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 88% select X against a Low label opponent, while 74% select X against a High label opponent.

These results are 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.

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 the large majority of subjects choose X even for the low payoff version, which goes against risk dominance.[7] 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

[7] Observe that assuming the utility function of money to be concave would not explain this result either, under dominance.

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 Additional games: Asymmetric Matching Pennies

In the Asymmetric Matching Pennies (AMP) game, we examine row and column players separately. Considering first the row players, we find that 44% of the High label players select X against another High label player, while 47% select X against a Low label player, in the low payoff version of the game. For the high payoff version of the game, 62% of the High label players select X, against both labels. For the Low label row players, we find that 52% play X against another Low label player. Against a High label player, this increases to 60%. For the high payoff specification, 62% of Low label row players select X against either label.

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.

For the column players, we find that High label column players in the low payoff specification choose Z with 67% frequency against both Low and High label players. This percentage increases to 79% and 73% against a Low, resp. High, label opponent. They 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 44% choose Z against another Low label player in the low payoff specification. This percentage increases to 56% against a High label opponent. For the high payoff specification, 72% choose Z against a Low label opponent and 67% against a High label opponent. Note that the model implies that their 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 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.

6 Conclusion

We have introduced in this paper a mechanism under which coordination can be achieved in isolated settings, without communication and without a notion of focality or shared culture. This mechanism relies instead on common agreement in a difference in cognitive sophistication between the players. Moreover, where coordination is expected to occur in our model depends on the payoff structure of the game. Specifically, we predict that coordination is more likely to be achieved on the preferred equilibrium of the player who is perceived to be less cognitively sophisticated in the canonical BOS game, but not in the modified BOS game we propose. We then test our predictions in a laboratory experiment. We find that the empirical findings support our model, in that the results are highly consistent with our predictions in the canonical BOS game. We also show that the results of the modified BOS game reject alternative mechanisms that may seem plausible, specifically a notion of a 'first-mover advantage' for the player seen as less sophisticated. Lastly, results for the additional games we consider, namely stag hunt and the asymmetric matching pennies, are also consistent with our model. Taken jointly, the experimental results strongly support the mechanism discussed in this paper.

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.

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.

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.

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.

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," .

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

Thomson, K. S. and D. M. Oppenheimer (2016): "Investigating an alternate form of the cognitive reflection test," *Judgment and Decision making*, 11, 99.

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