# Coordination and Sophistication*

Larbi Alaoui[†] Katharina A. Janezic[‡] Antonio Penta[§]

July 2025

**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 a variation of the Battle of the Sexes that we construct, for instance, the strategic advantage is reversed. We show that subjects' behavior in a laboratory experiment is consistent with the predictions of our model, and present evidence against alternative coordination mechanisms.

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

**JEL Codes:** C72; C91; C92; D80; D91.

## 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

*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 and BSE. E-mail: larbi.alaoui@upf.edu

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

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

view concerning which points are focal. In practice, however, and especially when agents face novel strategic situations, the conditions for shared focality to exist may not be met. In such cases, 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.

Although introspection might perhaps suggest that coordination should be possible at least in some such situations, up to date there is no mechanism to explain whether or under what conditions this might be achieved on purely *eductive* grounds. In this paper we show that in fact, even in the absence of focal points, coordination would be the outcome of a broad class of introspective reasoning processes, provided that two key conditions are met: first, that players' reasoning responds to incentives (cf. Alaoui and Penta 2016, 2022); second, that they view each other as having *different* cognitive abilities, and that they agree on their relative position. We also provide an experimental test of the theory and find that subjects' behavior is in line with the theoretical predictions.

As an example, consider two investors facing two new investment opportunities, on start-up $A$ or $B$. As is often the case in these situations, the two investors share a coordination motive (their returns are higher if they invest in the same asset, or projects only succeed if they attract both investors, etc.), but may also differ on which of the two alternatives they prefer to coordinate on (sources of such disagreement may be asymmetric information, heterogeneous beliefs, different portfolio holdings, etc.). Hence, the situation is akin to the BoS in Fig. 1. The two investors may be experienced or not, having played similar games in the past. This notwithstanding, if no communication is possible, and if neither investment opportunity is focal at the moment of their decisions, it is not obvious how they would manage to coordinate, if at all, or on which asset.

Now suppose that the investors have beliefs about 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 investors view each other to be of similar sophistication; in case (ii), they (commonly) believe that one is of higher sophistication than the other. For instance, investors may have various degrees of experience of similar strategic situations, and the view of sophistication may be based on that (then, case (i) materializes if investors face someone they regard as having similar experience, and case (ii) if they agree that one investor is much more experienced than the other).

In our model, these beliefs have clear implications on the likelihood of 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. This dichotomy has been extensively studied also from the viewpoint of general equilibrium theory (see Guesnerie; Guesnerie 2001; 2005 and references therein).

| | *Bach* | *Stravinsky* |
|---|---|---|
| *Bach* | $r, 50$ | $0, 0$ |
| *Stravinsky* | $0, 0$ | $50, r$ |

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

and on where the agents will coordinate. Specifically, we predict higher rates of coordination in case (ii) than in case (i). Furthermore, the model predicts that, in this game, players are more likely to coordinate on the equilibrium that is most favorable to the player who is believed to be *less* sophisticated. Hence, being perceived to be *more* sophisticated here is a disadvantage (as we discuss below, however, the opposite is true in other games). Moreover, and in contrast with what one might expect, payoff transformations that exacerbate the disagreement between players (while preserving the symmetry of the game, e.g., increasing $r$ in Figure 1), do not reduce coordination. In fact, it may favor its occurrence. These results bring to light a separate dimension of coordination, compared to the logic of coordination based on a shared culture and focality, which is typically associated with some form of homogeneity among players (cf. Kets and Sandroni 2019, 2021, Kets, Kager and Sandroni 2022, and Kets 2022). In the absence of focal points, it is agents' perceived *heterogeneity* that facilitates coordination.

The logic underlying our main results for the BoS game is the following. As players try to understand the game and their opponent's reasoning, they go through a sequence of actions that they consider playing (what we call 'path of reasoning'), until they stop, when the incentives to reason no longer compensate the extra cognitive costs. Then, if they think that the opponent has stopped earlier (e.g., because they have higher costs), they try to understand where they may have stopped, and choose optimally given their beliefs. Hence, the action that a player chooses is a function of both his own reasoning and his belief about that of his opponent.[2] In the BoS game, these cost-benefit criteria (which are derived from the axiomatization in Alaoui and Penta 2022) imply that if a player's favorite equilibrium is sufficiently preferred over the other one (i.e., if $r$ is high enough in Fig. 1), and if the path of reasoning is such that thinking more never loses the potential to change the player's action (what we call *responsive*), then it is more likely that a player's reasoning will stop at the action associated with his favorite equilibrium (this is true in the BoS, but not in other games). Now consider again the cases above, in which the players commonly believe that they are (i) of similar cognitive sophistication

[2] Our model 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). 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 et al. (2022). Halevy et al. (2021) is also related in that they design an experiment on strategic reasoning with more sophisticated opponents.

or (ii) of different sophistication. In the first case, the players play according to their own understanding of the situation, and may miscoordinate (unless their paths of reasoning are fortuitously aligned). In fact, as their disagreement over the two equilibria increases, they become *more likely* to miscoordinate. In the second case, instead, the player perceived to be less sophisticated plays according to his own understanding of the situation. But the more sophisticated player believes he has gone deeper than his opponent, and hence does not play according to where his own reasoning has stopped, but rather according to where he believes the opponent has stopped. 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. All is required is that they commonly agree on their relative sophistication. Also, our model is essentially unrestricted on the form of reasoning: 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 the reasoning path has to be responsive, which is a natural way of formalizing the inherent strategic uncertainty that arises in the absence of focal points.[3]

After introducing our model and the theoretical results, we present our experimental test of the theory. 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, for both a low and a high payoff $r$. In line with our predictions, we find that, for the BoS game: (i) high label subjects concede more against low than against high; (ii) this effect is more pronounced when $r$ is higher; (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. These results confirm our theoretical insights, including that in our setting it is heterogeneity rather than homogeneity that leads to increased coordination.[4]

The theoretical argument provided above might seem 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 introduce another coordination game in our experiment, which we refer to as the *reverse Strategic Advantage (RevSA)* game (see Figure 2). In this game, our model

[3]As we explain in footnote 6, in the model of Kets and Sandroni (2019, 2021) and Kets et al. (2022) behavior is generated by a reasoning process that stabilizes. Hence our focus on *responsive* paths of reasoning formally captures the sense in which our model is complementary to theirs.

[4]This may seem to contrast with the literature on in-group vs out-group behavior, in which shared culture can favor coordination via a common focal point. But since we focus on the occurrence of coordination *absent* a focal point, the role of shared culture is effectively turned off, while our mechanism remains present provided that there is agreement on relative cognitive sophistication.

| | *Bach* | *Stravinsky* |
|---|---|---|
| *Bach* | $130, 130$ | $230, r$ |
| *Stravinsky* | $r, 230$ | $170, 170$ |

Figure 2: A reverse Strategic Advantage Game, with $r \in [190, 220]$.

delivers the opposite prediction to the one obtained from the 'first mover' argument: here, the *high* type has the strategic advantage. A second natural conjecture is that the *asymmetry* in players' labels itself helps achieve more coordination. With this view, however, it would be difficult to explain how coordination favors the low label subjects in the BoS, and the high label in the RevSA game, as predicted by our theory. We find that the experimental 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, which have been included to assess the viability of some alternative mechanisms, such as risk dominance, that may guide subjects' choices in these games and to check whether the basic logic of the model also holds in non-coordination games (these are discussed in the Appendix). Put together, this set of results shows the empirical relevance of the mechanism introduced by our model.

The rest of the paper is organized as follows: Section 2 sets out the theoretical model and Section 3 contains the main theoretical results underlying the experiment. Section 4 presents the experimental design, the predictions, and results for the BoS game. Section 5 discusses competing explanations and tests thereof. 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 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 process, and that it is "as if" they perform a cost-benefit analysis in deciding whether or not to take one additional step of reasoning. That is, it is as if at a given step, players trade off the cognitive cost, which represents the difficulty of thinking they are currently experiencing, with some notion of *value of reasoning*, which is related to the game's payoffs. We take the cost and benefit functions to be incremental and myopic. The approach is as-if in the sense that we do not assume that this procedure is due to a deliberate, conscious calculation. Rather, to the extent that players' reasoning satisfies certain regularities, from the viewpoint of an external analyst it can be modelled as such. This is shown axiomatically in Alaoui and Penta (2022), for general stepwise reasoning processes.

## 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 $a_j \in A_j$.[5]

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. We note that the predictions of the model that we will analyze apply to a broad class of paths of reasoning $\{(a_i^{i,k}, a_j^{i,k})\}_{k\in\mathbb{N}}$, such as:

**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 $a_j^{i,k} \in BR_j(a_i^{i,k})$ 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.

**2. Level-$k$ Reasoning:** Another natural form of reasoning is level-$k$, introduced by Nagel (1995) (see also Crawford et al. 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

[5] 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.

action of an opponent of level ($k$-1). Formally, for each $k = 1, 2, ...$, this form of reasoning is such that $a_j^{i,k} = a_j^i (k-1)$ and $a_l^i (k) = BR\left(a_{-l}^i (k-1)\right)$ for each $l \in \{1, 2\}$, where $a^i (0) = \left(a_1^i (0), a_2^i (0)\right)$ is an arbitrary level-0 *anchor*. Note 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 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 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 $W_i$.

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 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 $\bar{k} = 0$, 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' $\bar{k} > 0$.[6] This property is best thought of as one way to capture a situation in which, if players could potentially reason indefinitely (as in Kets and Sandroni 2019, 2021, in which $k$ is infinite), they would potentially never stop questioning their earlier conclusions. In

[6]For instance, Kets and Sandroni (2019, 2021)'s *introspective equilibrium* (see also Kets et al. 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 our theoretical 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 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.[7]

As mentioned above, 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 depends on the game's payoffs, such as the $r$ parameter in the BoS game, and will be discussed shortly. 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 their 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 $c' \in C^+(c)$.*

[7] The theoretical analysis in this section focuses on individuals with *responsive* paths of reasoning. However, the predictions derived for the treatments in the experiment *do not* require that *all* individuals feature responsive paths of reasoning, but only a fraction. That is because, for individuals with absorbing paths of reasoning, their choice would either be affected by our treatments in the same way as those with responsive paths (that is, if the threshold $\bar{k}$ beyond which their path stabilizes has not been reached), or it would not be affected at all.

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.[8] 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.

In deciding whether to stop reasoning or not it is *as if* players have expectations about what action of the opponent they would learn at the next step of reasoning, and they anticipate that they would best respond to it. Then, they calculate the value of reasoning as the expected gain of switching from the current action, $a_i^{i,k-1}$, to such a best-response to what they might learn. So, for instance, they would attach a value of zero (and hence stop reasoning) if they expected their current conjecture $a_j^{i,k-1}$ to also be confirmed at the next step; but it may be positive otherwise. This formulation is consistent with the axiomatic foundation of Alaoui and Penta (2022). In the following, we maintain the most stringent parametrization within this class, where it is as if the agent assigns probability one that the next step of reasoning will yield the action of the opponent which maximizes the opportunity cost of stopping. Hence, we assume the following functional form:

$$v_i(k) = \max_{a_j \in A_j} u_i(BR_i(a_j), a_j) - u_i(a_i^{i,k-1}, a_j). \tag{3}$$

Less extreme forms of the value of reasoning, which for instance consider non-degenerate distributions over the opponent's actions that player $i$ may expect to learn (cf. Alaoui and Penta 2022), would not affect our main results. Hence, we use the *maximum gain* (or *maximum regret*) representation above because it 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.

An implicit assumption in the formulation above is the idea that transformations of the payoff functions do not affect the way in which individuals reason about the game

[8]This is because each step of reasoning adds to the set of hypothetical conjectures that a player is able to formulate about the opponent. Since players in our model know how to best respond if they believe their opponent has stopped at a lower level, such a set never shrinks, and hence the cost of the next step of reasoning involves both keeping the previous steps in the working memory as well as adding a new one.

(namely, their *path* and *cost* of reasoning), but only their incentives and hence possibly the *depth* of their reasoning. Since the payoff transformations that we focus on (such as varying the $r > 1$ parameter in the BoS game) do not change the fundamental structure of the game, this is a very weak assumption.[9] But while we assume that the path of reasoning is not affected by varying the $r$ parameter in the BoS game, we do not assume that an individual's path of reasoning is the same across different classes of games.

### 2.3 Beliefs about Others' Reasoning and Choice

The cognitive bound $\hat{k}_i$ describes the thought process of the agent, but his behavior also depends on his beliefs about his opponent, and particularly about 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.[10] 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 $j$'s 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:

$$\hat{k}_j^i = \min \left\{ \hat{k}_i, \mathcal{K}(c_j^i, v_j^i) \right\}. \tag{4}$$

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 $a_i = a_i^{i,\hat{k}_j^i}$, and hence we also refer to $k_i = \hat{k}_j^i$ as player $i$'s *behavioral level.*

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 to the player's own maximum cognitive bound: in the latter case, the cognitive bound is *binding* in the sense

[9] This property, for instance, need not hold for payoff transformations that change the nature of the game (e.g., turning a BoS into a Matching Pennies game). These ideas have been formalized in the axiomatic foundation of Alaoui and Penta (2022), with the notion of *cognitive equivalence.* First, two games are *cognitively equivalent* if the decision maker approaches them with the same reasoning process. Then, the representation theorems ensure that two cognitively equivalent games are associated with the same costs and path of reasoning, and only differ in the value of reasoning. Thus, it is meaningful to perform comparative statics based on the value of reasoning only within, but not across, equivalence classes.

[10] 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.

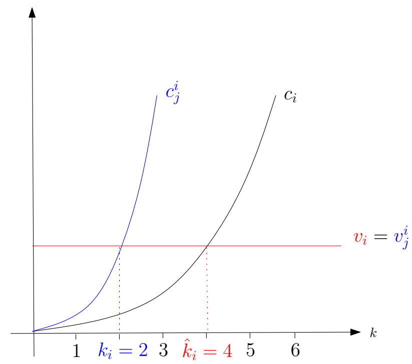



Figure 4: An example in which the cost $c_i$ and value $v_i$ are such that $i$'s cognitive bound $\hat{k}_i = \mathcal{K}(c_i, v_i) = 4$. The behavioral level $k_i$ is equal to 2 or 4 depending on whether the opponent is believed to be *less* or *more sophisticated* (respectively, on the left and on 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 $j$ 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 ($\hat{k}_i = k_i = 4$).

that the player's beliefs about the number of steps undertaken by his opponent are limited by his own cognitive bound. For the same reason, the following also holds:

**Remark 2** *If the path of reasoning is* absorbing*, then reasoning has ultimately no impact on what is learned past the threshold $\bar{k}_i$ where the path stops changing. But until such $\bar{k}_i$ is reached, reasoning has exactly the same effect as in the responsive case. Hence, in our model, when beliefs or payoffs change, the choice of an individual with an* absorbing *path of reasoning either does not change (that is, if he is already past his $\bar{k}_i$), or it changes in exactly the same way as it would for an individual with a* responsive *path of reasoning.*

We note that our formulation presumes that a player's value of reasoning does not depend on the opponent's payoff function, costs of reasoning, or his beliefs, but only on the player's own payoff function and current action (see Alaoui and Penta 2022 for an axiomatic foundation). That is, his *cognitive bound* (eq. (2)) does not depend on his beliefs about the opponent, of any order. Nonetheless, the *behavioral level* defined in eq. (4) does. Hence, while the player's cognitive bound (i.e., what he understands about the problem) is independent of his beliefs, behavior in our model may accommodate rich effects of players' first- and higher-order beliefs. For an experimental investigation of both, see Alaoui and Penta (2016) and Alaoui et al. (2020).

### 2.4 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.

In the next section we 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 (i) 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), and if (ii) both players agree on their relative sophistication.

# 3 Endogenous Coordination in the BoS game

## 3.1 Theoretical Results

In this section we present the main theoretical results that underlie our experiment. 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 $v_1(k) = \max\{1-1, r-0\} = r$ . 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 $r$. Hence, if $r$ increases, the maximum gain increases at steps where $a_1^{1,k-1} = W_1$, but it remains at 1 at steps in which $a_1^{1,k-1} = B_1$. 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 that is *responsive*. 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 $a_1^{1,\hat{k}_1} = W_1$, 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 $r$

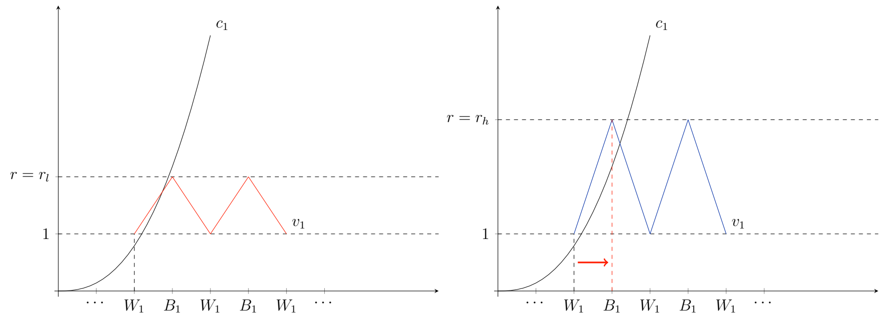



Figure 5: Low-payoff cognitive bound such that $a_1^{1,\hat{k}_1} = W_1$.

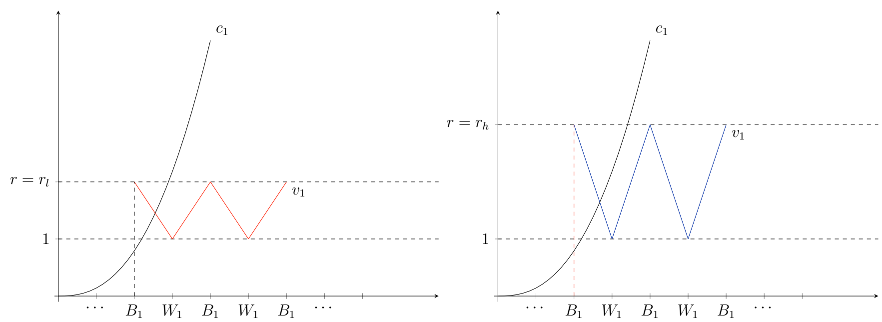



Figure 6: Low-payoff cognitive bound such that $a_1^{1,\hat{k}_1} = B_1$.

has associated $a_1^{1,\hat{k}_1} = B_1$, then an increase in $r$ has no effect. Whereas the actual step $\hat{k}_1$ at which the agent stops may vary in the two cases, in either case it would be such that $a_1^{1,\hat{k}_1} = B_1$ for high enough $r$.[11]

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

[11] We note that the same logic would also apply to less extreme forms for the value of reasoning function, where it is *as if* player $i$ has beliefs about what he could learn about the opponent's action that are not concentrated on the $a_j$ that maximizes the opportunity cost of playing the current action, $a_i^{k-1}$. In $2 \times 2$ games, any such non-degenerate beliefs would induce a scaled-down version of the 'maximum gain' value of reasoning, which would affect the level of the $\bar{r}_i$ threshold in the statement of Lemma 1, but not its existence. The main advantage of the maximum gain representation is that it has no free parameter.

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 $r \geq 1$, $\mathcal{K}(v_i, c_i) < \mathcal{K}(v_j^i, c_j^i)$ (resp., $c_j^i \in C^-(c_i)$ and $\mathcal{K}(v_i, c_i) > \mathcal{K}(v_j^i, c_j^i)$).

**Example 1** *As an example of this notion, consider a parametric model, in which the cost functions are assumed to be linear in $k$ and the reasoning process for the two players in the BoS in Fig. 3 has a cyclicality of* 1 *(i.e., for both $i$, $\pi_i \in \Pi$ only if $\pi_i$ such that, for any $k$, $a_i^{i,k} = B_i$ if and only if $a_i^{i,k+1} = W_i$; e.g., Fig. 5 and 6). Then, for $c_i(k) = k/\alpha_i$, and $c_j^i(k) = k/\alpha_j^i$, we have that $c_j^i \in C^-(c_i)$ for all $\pi_i \in \Pi$ if $\alpha_j^i < \lfloor \alpha_i \rfloor - 2$, where $\lfloor \alpha_i \rfloor := max\{n \in \mathbb{N} : n \leq \alpha_i\}$.*[12] *Besides illustrating the concept, this example obviously provides sufficient conditions for the notion above to hold within a parametric version of our model. In order to highlight the core theoretical idea, however, the results that follow, are stated for the general model and the general notion of the sets $C^+(c_i)$ and $C^-(c_i)$.*

Given this, if $i$ believes that $j$ is strictly 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 strictly 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

[12] More generally, in a BoS with 0 non-equilibrium payoffs where $R > 0$ and $R + r > R$ are, respectively, the lowest and highest equilibrium payoffs, then $c_j^i \in C^-(c_i)$ for all $\pi_i \in \Pi$ if $\alpha_j^i < \lfloor \alpha_i \rfloor - 2/R$. Note that the argument above does not require that any particular knowledge of the starting point of the other player, or that the agents' paths be aligned. All the implications follow from a 'within-player' logic.

agree about their relative ranking. A noteworthy implication of Proposition 2 is that, conditional on being in a 'heterogeneous matching', it is the relatively *less* sophisticated player who has a strategic advantage. As mentioned in the introduction, however, this is a specific feature of this game, and does not hold in general. For instance, in the 'Reverse Strategic Advantage' game that we present in Fig. 2, it is the *more* sophisticated player that gets a strategic advantage. This raises interesting questions about individuals' incentives to be perceived as more or less sophisticated as a function of the setting. We return to these points in the conclusions.

We now turn to the case where there is no commonly agreed ranking about players' relative sophistication. In that case, it may be that both players think that the other is *less* (respectively, *more*) sophisticated than they are themselves. In those cases, miscoordination would ensue: if both players believe that they are the relatively *less* sophisticated player, then they would both be driven by their cognitive bound, and hence play the action associated with their most preferred equilibrium in the BoS game. The opposite is true if they both think the opponent is less sophisticated: in that case, they would both concede, and play the action associated with the equilibrium most preferred by their opponent.

**Proposition 3 (Miscoordination in the BoS)** *Under the maintained assumptions, in the BoS, if both players' paths of reasoning are* responsive *and if they* do not agree on their relative sophistication*, then there exists $\bar{r}_i$ such that, for all $r > \bar{r}_i$, players play $a = (B_i, B_j)$ if they both regard the opponent as more sophisticated, and $a = (W_i, W_j)$ if they both regard the opponent as less sophisticated.*

The fact that Propositions 1 and 2 obtain 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 equilibrium 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 distills the essence 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.

### 3.2 Extensions and Applications

We note that the model can be applied to general settings, but the results on coordination would depend on the specific situation. The analysis of 2x2 games, however, is especially useful to distill the logic of the model and the fundamental source of the coordination result (namely, agreement about players' heterogeneous abilities, and responsiveness of the path of reasoning). While extending the analysis to other games is left to future research, the fundamental logic that we are highlighting is a useful guideline. For instance, for an application to a 'mass action' game, consider two populations of agents, taking the roles of the row and column player, respectively. Each agent must choose between two actions (e.g., strike or not, attack one currency or another, attack or not, invest or not, etc.).[13] To fix ideas, suppose that the two populations of agents are members of different organizations, deciding which of two governments/firms to attack/protest against. Within-organization coordination is made possible by communication, by a common leadership, or by a deliberation procedure internal to the organization. But the coordination problem remains between the two separate organizations. In this setting, both the logic and the results above apply equally well: if members of the two organizations commonly agree that the members of one organization have lower costs of reasoning than the other, then coordination would arise for sufficiently high payoff parameters $r$. Such coordination would occur on the equilibrium most favorable to the 'high cost' organization if payoffs are akin to a BoS game, on the other equilibrium if they are closer to a RevSA game (cf. Fig. 2), on the efficient equilibrium if payoffs are as in Stag Hunt, etc. If, in contrast, the two organizations failed to agree that the members of one organization are more sophisticated than those of the other organization, then coordination need not be achieved, and it would not be achieved if $r$ is high enough.

The results in the paper would also extend to games with more than two actions, except that the coordination result would require a more careful (and somewhat more cumbersome) formulation of the notion of *responsive paths*. A simple requirement would be that the path of reasoning 'visits infinitely often all of the actions associated with all the equilibria in the game'. Under this condition, the analysis above could be extended to such settings. Alternatively, this assumption can be weakened by making stronger assumptions on players' beliefs, paths of reasoning, or cognitive costs, and for suitably defined payoff transformations that increase players' incentives to reason. For instance, in coordination games with Pareto ranked equilibria, the logic of our results would apply, regardless of the number of actions, to support efficient coordination, as long as players' reasoning paths never settle on ruling out the efficient equilibrium actions.

[13] We note that it is very frequent in the literature to cast "mass action games" as 2x2 games in this fashion, re-interpreting mixed actions in the game in terms of fraction of the populations of row and column players taking one action or the other, often using precisely games such as BoS, Stag Hunt, etc. See, for instance, Morris and Shin 2003 on global games and applications, and references therein.

# 4 The Experiment

## 4.1 Experimental design and logistics

The experiment is designed to test whether behavior in the BoS game is in line with our hypotheses, which are derived from the propositions in the previous section and are stated below. It also includes other games, which will be discussed in the next section and in the appendix, that allow us to test whether the results can be explained by alternative theories instead. The data used in the experiment is available at Alaoui et al. (forthcoming).

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 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 Raven test at the end of the experiment, and vice versa). Results for the comparison between the two tests are given in Appendix A.3.5.

Subjects are then separated into three groups, depending on whether their scores were High, Moderate, or Low. Cutoffs were predetermined and based on the distributions of test scores obtained in Alaoui and Penta (2016), Alaoui, Janezic and Penta (2020) and subsequent pilots. The 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. The High and Low groups are informed of their labels, the Moderate group is not. Subjects are only informed of their label, not of the scoring rule used to classify the subjects. For the main experiment, we use only the High and Low groups, in order 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 is large enough to generate different depths of reasoning (p. 15). Since only the High and Low groups are relevant for our purposes, all discussions below refer to these groups. The Moderate group plays an unlabeled treatment, as documented in Appendix A.3.6.

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:

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

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 $X$ ($Z$) interchangeably (and, respectively, $W_i$ and $Y$ ($W$)). The experiment employs a within-subject design, with every subject playing each of the following four versions of the game, in the role of the *row player*, without feedback and with random anonymous matching at every round:[14]

- $\mathbf{BoS_{same}}$: The BoS game is played against someone with the same label, for the smaller reward $r = 51$.
- $\mathbf{BoS_{other}}$: The BoS game is played against someone with the other label, for the smaller reward $r = 51$.
- $\mathbf{BoS^{+}_{same}}$: The BoS game is played against someone with the same label, for the greater reward $r = 70$.
- $\mathbf{BoS^{+}_{other}}$: The BoS game is played against someone with the other label, for greater reward $r = 70$.

In addition to the BoS game, subjects also play the *reverse Strategic Advantage* game, as well as a Stag Hunt and an Asymmetric Matching Pennies game. As with the BoS, they each play four versions of these three games: against an opponent with the same label, against someone with the other label, both for the low and the high payoff versions of the games. These games are included to assess the viability of some alternative mechanisms that may guide subjects' choices in these games, and will be discussed in Sections 5, A.3.3 and A.3.4. 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. See Figure 18 in Appendix A.4.1 for an illustration of the design.

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 the Moderate group and participated in the unlabeled treatment.

[14] In Appendix A.5, we provide a glossary of the terminology used to refer to the different treatments.

## 4.2 Experimental Hypotheses for the BoS game

The theoretical results in Sections 2 and 3 refer to a setting where all beliefs are degenerate. This was done to best highlight the core logic of the model, and the role that agreement on relative sophistication plays for achieving equilibrium coordination. Of course, beliefs need not be degenerate in practice, and various forms of noise may intervene in the core logic that we distilled above. Here, we discuss the different variations of the baseline model that are applicable, and how they will be accounted for in our experimental hypotheses.

There are several kinds of probabilistic beliefs and stochasticity one could take into account: (1) non-degenerate beliefs over the costs of the opponent; (2) stochasticity of the reasoning process; and (3) costs of reasoning or beliefs over the opponent's costs that may vary stochastically throughout the experiment, and may therefore lead to stochasticity in individual behavior, and heterogeneity across subjects.

The first two variations would not affect any of the model's predictions in a significant way. For the first point, as discussed in Footnote 10 and Section 3.1, as long as the supports of the beliefs over the opponent's costs are all on one side of one's own cost of reasoning, the fact that they are non-degenerate would only affect the critical thresholds that determine the response to the change in the incentive to reason (the $\bar{r}$ in the proposition), but not the results they trigger. The second kind of variation instead would indeed add a layer of stochasticity to our results, but again it would not remove the main mechanism, nor affect the comparative statics, since the results above do not rely on a specific type of reasoning process. The third point instead is more subtle, and we explicitly account for it in the experimental hypotheses that follow.

To account for the possibility that not all agents of each label (Low or High) in our experiment share the same belief over the relative sophistication of their opponents, we will assume that only a portion of subjects believes that they are more sophisticated than the opponent that they are faced with, as a function of their own label and of their opponent's label. In particular, we will assume that a higher fraction of subjects (of both labels) believes that those with a High label are strictly more sophisticated than those with a Low label, and that less than half view their opponents as strictly less sophisticated than themselves when facing their own label. Furthermore, to take into account potential stochasticity throughout the experiment, these fractions will be assumed to be drawn randomly from round to round, for each agent.

Formally, let $L$ denote the label of subjects who were classified as Low on the test, and $H$ the label of those who were classified as High, and let the generic label be denoted as $l \in \{L, H\}$. Recall that subjects are informed of their own and their opponents' labels. Subjects in the experiment are heterogeneous, and as in our model, within each treatment $[t]$ (or *round* in the experiment; that is, for the BoS game, $[t] \in \{BoS_{same}, BoS_{other}, BoS^{+}_{same}, BoS^{+}_{other}\}$) each subject is identified by a label $l \in \{L, H\}$ and a tuple $i_{[t]} = (c_i, c^i_j, \pi_i) \in C_i \times C^i_j \times \Pi$, which we refer to as *type*, where we let $C = C_i = C^i_j$ denote the set of costs of reasoning functions that satisfy the conditions in

Assumption 1, and $\Pi$ denotes the set of paths of reasoning. We accommodate heterogeneity among agents, and let $\rho^{[t,l]} \in \Delta(C_i \times C_j^i \times \Pi)$ denote the distribution of types in the population of subjects with label $l$ in treatment $[t]$, and we will let $\rho_X^{[t,l]} := \text{marg}_X \rho^{[t,l]}$. By *not* assuming that $\rho_X^{[t,l]} = \rho_X^{[t',l]}$ for any $t, t'$ and any $X$, we accommodate stochasticity throughout the experiment. Finally, we let $\Pi^* \subset \Pi$ denote the set of paths of reasoning that are *responsive*.

**Assumption 2 (Identification Assumptions)** *We maintain the following assumptions about the population distribution in each treatment* $[t]$*:*

1. *Paths of reasoning are* responsive *for at least some percentage of L and H subjects. That is, for both l,* $\rho_\Pi^{[t,l]}(\Pi^*) > 0$.

2. *When matched with someone of the* same *label, subjects have* heterogeneous beliefs, *with a fraction* $q_s^{[t]} \in [0, 1/2)$ *believing that they face a strictly less sophisticated opponent. That is, for both labels,* $\int_{c_i \in C_i} \rho^{[t,l]}(\{c_i\} \times C^-(c_i) \times \Pi) =: q_s^{[t]} \in [0, 1/2)$.

3. *When matched with someone of the* other *label, subjects have heterogeneous beliefs, but the fraction of subjects that within a match think that the H-label is strictly less sophisticated than the L-label is bounded above by* $q_s^{[t']}$*, where* $[t']$ *denotes the 'same label treatment' with the same payoff specification as* $[t]$*. That is, the fraction of L subjects that think the opponent is less sophisticated is in* $\left[0, q_s^{[t']}\right)$*, and the fraction of H subjects that think the opponent is less sophisticated is in* $\left(1 - q_s^{[t']}, 1\right]$*, or more formally,* $\int_{c_i \in C_i} \rho^{[t,L]}(\{c_i\} \times C^-(c_i) \times \Pi) =: q_o^{[t]} \in \left[0, q_s^{[t']}\right)$ *and* $\int_{c_i \in C_i} \rho^{[t,H]}(\{c_i\} \times C^-(c_i) \times \Pi) =: 1 - q_o^{[t]} \in \left(1 - q_s^{[t']}, 1\right]$.

Note that, if $q_s^{[t]} \in (0, 1/2)$, then the model accommodates heterogeneity in beliefs in all treatments, in the sense that some agents do not necessarily share the same ranking between their costs of reasoning and that of the subject they are matched with. If instead $q_s^{[t]} = 0$, then beliefs are homogeneous in the sense that all types think the opponent is strictly less sophisticated when they share the same label, and all subjects agree that $H$ labels are more sophisticated than $L$ labels in mixed labels matchings (note, however, that we still allow for heterogeneity of the specific cost of reasoning $c_i$ and beliefs $c_j^i$, as well as of the paths of reasoning). We will refer to the case $q_s^{[t]} = 0$ as the *common agreement* model. This special case is extreme, but it represents a useful benchmark nonetheless, since it matches exactly the explanation in the theoretical section above. The comparative statics that we derive in the following, however, which concern shifts of the distribution of actions across various treatments, and which form the basis of our experimental hypotheses, hold true for any $q_s^{[t]} \in [0, 1/2)$, under the maintained assumptions above.

Assumptions 2.2 and 2.3 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. As we explain below, strictly speaking, Assumption

2.1 is not required for the hypotheses that follow. Without this assumption, however, our model has no bite, since all our hypotheses are predicated on the assumption that at least *some* paths of reasoning are responsive. We *do not* require that *all* individuals have responsive paths of reasoning, because the choices of agents with non-responsive paths would either not be affected by the treatments in our experiment, or move in the same direction as those that we derived for responsive paths of reasoning (see Remark 2). Hence, the comparative statics that underlie the hypotheses below are driven by the results obtained for the latter paths of reasoning. Overall, these are very weak assumptions, particularly within the context of our experimental design.

For $l \in \{L, H\}$, let $p^l(\cdot)$ denote the percentage of subjects in group $l$ that play their own preferred action, where the argument of the function refers to the treatment. Assumption 2 and Propositions 1-3 directly imply the testable hypotheses that follow.

**Hypothesis 1 (Same to other label (opponent) comparison)**

1. $p^H(BoS_{same}) \geq p^H(BoS_{other})$ *and* $p^H(BoS^+_{same}) \geq p^H(BoS^+_{other})$*: 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. $p^L(BoS_{same}) \leq p^L(BoS_{other})$ *and* $p^L(BoS^+_{same}) \leq p^L(BoS^+_{other})$*: the percentage of Low subjects playing their own preferred action in the BoS game is* higher *when playing against subjects with the other label than against subjects with the same label, for both values of* $r$*.*

Note that, as stated, these hypotheses also hold without Assumption 2.1. This is because even if no subjects have a responsive path, then their behavior would either not change (if their reasoning is at the 'absorbing' part of the path, i.e., if the $k$ is weakly above their $\bar{k}$) or change in the direction predicted by Hypothesis 1 (if they are at the non-absorbing part of the path, if that part exists), both of which are allowed by the weak inequalities. But if it were the case that all the subjects had non-responsive paths, and further that they were at the absorbing part of their path, then the mechanism discussed here would never be switched on. Hence, we would observe no change even as payoffs are varied, which we turn to next. As can be seen from our findings, the experimental results are indeed consistent with Assumption 2.1, or at least that at least some percentage is at the non-absorbing part of their path.

In addition to Hypothesis 1, in which we consider behavior as the opponent's label is varied but the payoffs remain the same (for both values of $r$), the model's predictions also yield the following testable hypotheses, as $r$ is varied but the opponent is kept fixed:

**Hypothesis 2 (Low to high payoffs comparison, Heterogenous Matchings)**

1. $p^H(BoS_{other}) \geq p^H(BoS^+_{other})$: *in the BoS, the percentage of high subjects playing their own preferred action is weakly* decreasing *in r when playing the other label.*

2. $p^L(BoS_{other}) \leq p^L(BoS^+_{other})$: *in the BoS, the percentage of low subjects playing their own preferred action in the BoS game is weakly* increasing *in r when playing the other label.*

Under the added condition that the $q_s^{[t]}$ in Assumption 2 is low enough, the following extra Hypothesis follows:[15]

**Hypothesis 3 (Low to high payoffs comparison, Homogenous Matchings)**

1. $p^H(BoS_{same}) \leq p^H(BoS^+_{same})$: *in the BoS, the percentage of high subjects playing their own preferred action is weakly* increasing *in r when playing their own label.*

2. $p^L(BoS_{same}) \leq p^L(BoS^+_{same})$: *in the BoS, the percentage of low subjects playing their own preferred action is weakly* increasing *in r when playing their own label.*

Hypotheses 2 and 3 show more subtle implications of our model which would arguably be challenging to replicate with other mechanisms or alternative explanations. The reasoning behind this is best seen for the 'common agreement' case, when $q_s = 0$. In this case, the high types are playing according to their bound when playing their own label, and as $r$ increases, there is a higher percentage of high types whose bound will be at their own preferred action. In particular, subjects for whom $\bar{r}_i \in [51, 70]$, may switch their chosen action 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 subjects believes 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. By continuity, the same comparative statics remain true when $q_s$ is small enough, in which case both Hypothesis 2 and 3 follow. For larger $q_s$, instead, the comparative statics over the homogeneous label treatments are indeterminate without stronger assumptions on the path of reasoning, whereas those over the heterogeneous treatments remain true.

We test Hypotheses 1-3 using a paired Wilcoxon signed-rank test and the McNemar test (since they yield identical results, we only report results for the former). To account for worse performance of non-parametric tests in smaller samples, we also repeat the analyses with a t-test. Results are in line with those from the Wilcoxon and the McNemar

[15] Formally, there exists $\bar{q} \in [0, 1/2)$ s.t. Hypothesis 3 follows from Assumption 2 whenever $q_s^{[t]} \in [0, \bar{q}]$. Also note that, if *common agreement* were imposed everywhere (i.e., if $q_s^{[t]} = 0$), then the conditions in Hypothesis 1.2 would hold with equality.

tests. We also use panel regressions with individual fixed effects. All regressions share the following specification:

$$Y_{i,t} = \beta X_{i,t} + \alpha_i + u_{i,t}$$

where $Y_{i,t}$ is the dummy variable capturing whether an individual chooses their favorite equilibrium action, $X_{i,t}$ is the dummy variable of whether they are playing against an opponent with an $H$ label (for tests of Hypothesis 1), or a dummy of whether the game is of the high payoff version (for tests of Hypotheses 2 and 3), $\alpha_i$ are individual fixed effects and $u_{i,t}$ the error terms. For Hypothesis 1, we conduct the regressions twice for each subject label, one model for the low payoff games and one model for the high payoff games. To test Hypotheses 2 and 3, we also conduct the regression twice for each label, one model for games played against the $L$ label and one model for games played against the $H$ label. The regressions use standard errors clustered at the individual level. To check robustness against possible small sample effects, we repeat the regression analyses with jackknifed, as well as with bootstrapped, standard errors. The results are consistent across the three specifications. Tables for the bootstrapped and jackknifed regressions are given in Appendix A.1.

We note that while all of the hypotheses above are about individual behavior, in Appendix A.3.1 we also discuss whether there is increased coordination in the heterogeneous treatments on the preferred action profile of the $L$ label subjects.

### 4.3 Results of the BoS Game

In this section, we discuss the results relating to Hypotheses 1, 2 and 3. Recall that subjects always choose in the role of the row player. All of the reported analyses use the pooled sample of subjects, combining subjects who are classified based on our test with those classified using the Raven matrices test.

According to Hypothesis 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. Figure 8 shows the proportions 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 Figure 8). However, when they face a Low label player, this percentage drops to 34.91%. We compare the distribution of chosen actions using a Wilcoxon signed-rank test.[16] The p-value of the test statistic is 0.003. In addition, we conduct a fixed

[16]Note that whenever we give results for the Wilcoxon signed-rank test, we are referring to the paired version of the test. For ease of exposition, we will omit 'paired'. We provide the exact p-values, unless stated otherwise. As mentioned above, the McNemar test for binary variables gives identical p-values. The results of the t-tests are in line with the results from the Wilcoxon and the McNemar tests.

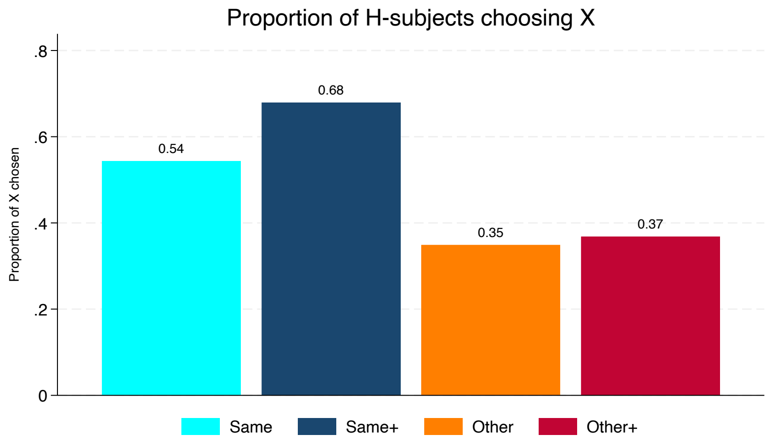



Figure 8: Results BoS Game - High Label Players: Proportion choosing X (their preferred action). *Same* (*Same+*) refers to the low (high) payoff version of the BoS game played against another player from the same label. *Other* (*Other+*) refers to the low (high) payoff version of the BoS game played against another player from the other label.

effects panel regression with the High players, restricting the sample to the low payoff version of the BoS, 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 (regression results are given in Table 1, Model (1)). These findings are all consistent with Hypothesis 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 p-value of the Wilcoxon signed-rank test statistic is less than 0.001. The regression coefficient is also significant at more than 0.1% (see Table 1, Model (2)). 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 Hypothesis 1.1.

We next turn to Hypothesis 1.2, which predicts that the Low label subjects would be less 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 Figure 9. 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 (Wilcoxon signed-rank test, p-value$=0.541$). We run equivalent panel regressions to the $H$ label discussed above. The regression results, in Table 2, Model (1), show that there is no significant effect of playing against a high label opponent.

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 (Model (2)) and the Wilcoxon signed-rank test statistic is also

| | Low payoff games | High payoff games |
|---|---|---|
| | (1) | (2) |
| | Choice of X | Choice of X |
| Opponent has H label | 0.204*** | 0.311*** |
| | (3.19) | (4.60) |
| Constant | 0.344*** | 0.368*** |
| | (10.94) | (10.87) |
| Observations | 209 | 212 |

*t* statistics in parentheses

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

Table 1: BoS Game: Panel (fixed effects) regression results for *H* label players.
Model (1) gives the results for the low payoff versions of the BoS game and (2) for the high payoff versions. Standard errors are clustered at the subject level.

| | Low payoff games | High payoff games |
|---|---|---|
| | (1) | (2) |
| | Choice of X | Choice of X |
| Opponent has H label | 0.0930 | 0 |
| | (0.81) | (0.00) |
| Constant | 0.535*** | 0.605*** |
| | (9.30) | (10.89) |
| Observations | 86 | 86 |

*t* statistics in parentheses

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

Table 2: BoS Game: Panel (fixed effects) regression results for *L* label players.
Model (1) gives the results for the low payoff versions of the BoS game and (2) for the high payoff versions. Standard errors are clustered at the subject level.

not significant (all have p-value= 1). The results for the Low label subjects are consistent with Hypothesis 1.2. The lack of significance for the low payoff result might be due to lower power, but goes in the direction suggested by Hypothesis 1.2. The finding of no change in the proportion of $X$ played in the high payoff games is also consistent with the hypothesis.[17]

Testing next Hypotheses 2.1 and 3.1, we examine whether it is the case that 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 against Low, as payoffs are increased. We indeed see from Figure 8 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 weakly statistically significant

[17] Recall that, if *common agreement* were imposed, then the model would imply no change in these treatments (cf. footnote 15).

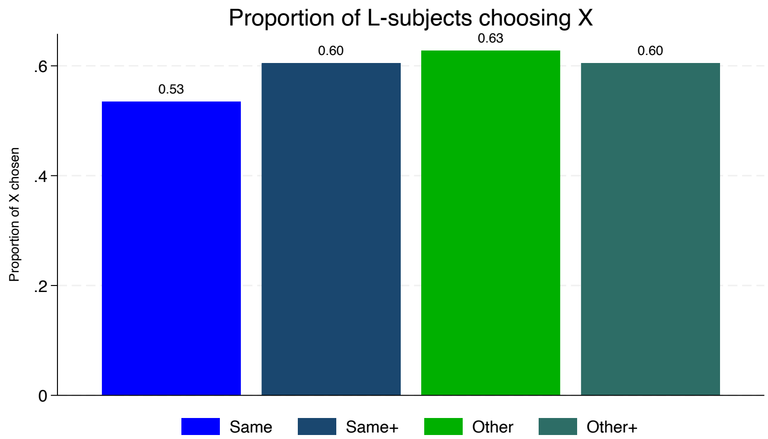



Figure 9: Results BoS Game - Low Label Players: Proportion choosing X (their preferred action). *Same* (*Same+*) refers to the low (high) payoff version of the BoS game played against another player from the same label. *Other* (*Other+*) refers to the low (high) payoff version of the BoS game played against another player from the other label.

at the 10% level (Wilcoxon signed-rank test, p-value= 0.092, and see the corresponding panel regression results given in Model (1) of Table 5 in the Appendix). When playing against Low opponents, the percentage playing $X$ increases from roughly 35% to around 37%, but this is not significant (Wilcoxon signed-rank test, p-value= 0.851). These results support Hypotheses 2.1 and 3.1.

Hypotheses 2.2 and 3.2 state 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 Figure 9 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 hypotheses. The percentage decreases against a High opponent from 62.79% to 60.47%, but these results are both not significant (Wilcoxon signed-rank test p-values= 0.648 and $\approx 1.000$, respectively. See as well Table 7 in the Appendix).

We also check whether our results above are driven by a few outliers (see Appendix A.3.2). We find that behavior is overall consistent with the model's predictions across the sample. Jointly, the findings are consistent with our hypotheses for the BoS game. Moreover, while the Hypotheses are all 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. Similarly, the fact that also Hypothesis 3 is corroborated is consistent with the fact that the fraction of subjects that, when matched with someone of the same label, believe they face a less sophisticated opponent, is sufficiently small.

Finally, note that Hypotheses 1-3 jointly imply higher coordination rates in the *other label* (High vs. Low labels) than in the *same label* (High vs. High and Low vs. Low) treatments, at least for sufficiently high incentives. We indeed find that coordination

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

Figure 10: A *reverse* Strategic Advantage Game, with $r \in [190, 220]$.

rates are higher in *other* than in *same* label treatments, and on the preferred action profile of the Low labels (see Appendix A.3.1).

# 5 Competing Explanations

There are at least two alternative theories that one might consider to explain the experimental results in the previous section. The first is the view that the increased coordination that we observe in the *other label* treatments (High vs. Low) 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 (FMA)*, 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 *reverse Strategic Advantage (RevSA)* game that we presented in the introduction, which we reproduce in Fig. 10 labeling players' actions $B_i$ and $W_i$ to denote, respectively, the action associated with the *'best'* and *'worst'* equilibrium for player $i$.

This game leads to 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, within the relevant parameter range of $r \in [190, 220]$, it is the $W_i$ action that is associated with the higher value of reasoning, and such a value is increasing with $r$.

Hence, in the *RevSA* game, the prediction of our model contrasts with the view that the low type receives a *first-mover advantage*: in this game, such a theory would predict that equilibrium coordination occurs on the profile 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.

Finally, we note that the latter discussion also addresses another possible view, according to which, perhaps due to fairness concerns, higher sophistication subjects may willingly 'give in' and allow less sophisticated opponents to obtain a higher payoff. If this view is aligned with the prediction of our model in the BoS game, it yields an opposite prediction for the *RevSA* game. Hence, these two games can also be used to discern our model from such an alternative 'fairness-based' explanation.

### 5.1 FMA vs Cost-Benefit: Testable Hypotheses

The game used in the experiment takes exactly the form given in Figure 10, 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 the 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 4.2, our model implies the following testable hypotheses in the *RevSA* game:[18]

**Hypothesis 4 (Cost-Benefit in the RevSA Game)** *Under our Cost-Benefit model, in the RevSA game the following holds:*

1. $p^H(RevSA^+_{same}) \leq p^H(RevSA^+_{other})$*: 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. $p^L(RevSA^+_{same}) \geq p^L(RevSA^+_{other})$*: the percentage of Low 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 same compared to with the other label.*

3. $p^L(RevSA_{other}) \geq p^L(RevSA^+_{other})$ *and* $p^H(RevSA_{other}) \leq p^H(RevSA^+_{other})$*: when playing against the other label, the percentage of subjects playing their own preferred action is weakly increasing in* $r$ *for the High sophistication subjects, and weakly decreasing for the Low sophistication subjects.*

4. $p^L(RevSA_{same}) \geq p^L(RevSA^+_{same})$ *and* $p^H(RevSA_{same}) \geq p^H(RevSA^+_{same})$*: when playing against an opponent with the same label, the percentage of subjects playing their own preferred action in the RevSA game is weakly decreasing in* $r$ *for both High and Low sophistication subjects.*

Note that these predictions are opposite to the view that the low type gets a first-mover advantage (FMA) under heterogeneous matching. Maintaining the same notation, the predictions of the FMA-model lead to the following testable hypotheses:

[18] Analogous to the discussions of Hypotheses 1-3 above, points 1-3 in Hypothesis 4 follow from Assumption 2, whereas point 4 follows under the extra assumption that $q_s^{[t]}$ is small enough (footnote 15). Here too, if common agreement were imposed ($q_s^{[t]} = 0$), the condition in point 2 would hold with equality.

**Hypothesis 5 (FMA in the RevSA Game)** *Under the FMA-view, in the RevSA game the following holds:*

1. $p^H(RevSA^+_{same}) > p^H(RevSA^+_{other})$
2. $p^L(RevSA^+_{same}) < p^L(RevSA^+_{other})$.

### 5.2 RevSA Game Results

Figure 11 displays the choices of the High label players in the four versions of the RevSA game. For the High label group, more than 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%. As with the BoS game, we conduct a paired 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 0.014. We also conduct panel regressions (Table 9 in the Appendix) to assess whether changing the opponent has a significant effect on the action choice and find that the effect is significant at the 1% level. 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.

To examine Hypothesis 4.1, we look at the high payoff version of the game. Here, 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. This goes against the hypothesized direction, but the increase is not statistically significant (p-value of 0.720 for the Wilcoxon signed-rank test, see also regression results in Table 9). Importantly, in both the low and the high payoff games, there is no evidence of High label players conceding to the Low label players, which goes against the first-mover advantage argument.

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 Figure 12). This difference is not statistically significant (p-value= 0.845, 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. While the direction of this difference is in line with Hypothesis 4.2, it is not statistically significant (p-value= 0.804, Wilcoxon signed-rank test). We also note that this change is in the direction opposite to that of the first-mover advantage theory.

To examine Hypotheses 4.3 and 4.4, we study how observed behavior changes when the opponent is held fixed and payoffs increase. To do so we again conduct panel regressions (see Tables 11 and 12 in the Appendix) as well as Wilcoxon signed-rank tests.

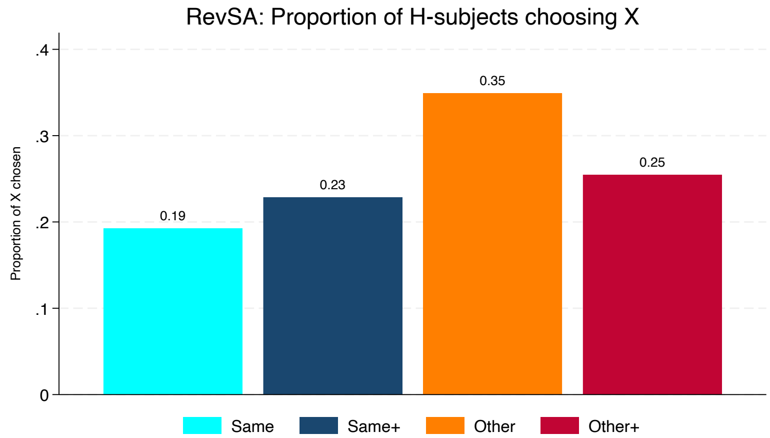



Figure 11: Results RevSA Game - High Label Players: Proportion choosing $X$ (their preferred action). *Same* (*Same+*) refers to the low (high) payoff version of the RevSA game played against another player from the same label. *Other* (*Other+*) refers to the low (high) payoff version of the RevSA game played against another player from the other label.

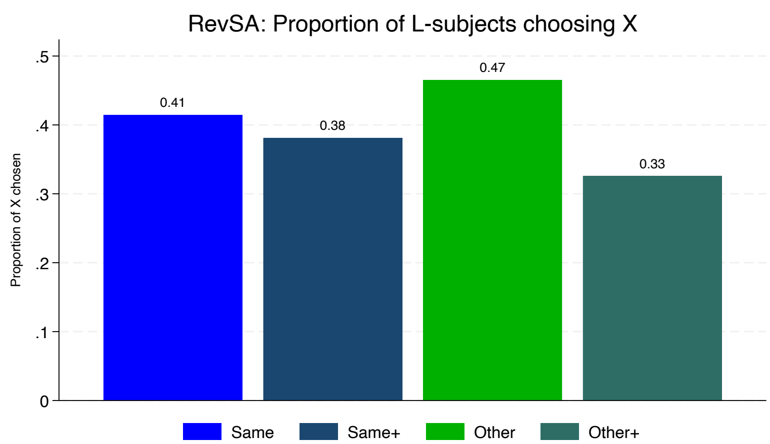



Figure 12: Results RevSA Game - Low Label Players: Proportion choosing $X$ (their preferred action). *Same* (*Same+*) refers to the low (high) payoff version of the RevSA game played against another player from the same label. *Other* (*Other+*) refers to the low (high) payoff version of the RevSA game played against another player from the other label.

Turning first to Hypothesis 4.3, for the games played by $H$ against $L$, around 35% of $H$ subjects in the low payoff game and around 25% in the high payoff game choose their preferred action. The direction of the effect is contrary to Hypothesis 4.3. The effect is weakly significant at the 10% level in the regression and has a p-value of 0.110 in the Wilcoxon signed-rank test. In the treatments in which the $L$ labels play against $H$ opponents, around 47% of $L$ subjects choose their preferred action in the low payoff game and close to 33% choose the same action in the high payoff game. The direction of the effect is as hypothesized, but not statistically significant (p-value=0.238, Wilcoxon signed-rank test).

Turning next to Hypothesis 4.4, when $H$ labels play against other $H$ label players, around 19% choose their preferred action in the low payoff version and around 23% do so in the high payoff game. The direction of the effect goes in the opposite direction to that predicted but is insignificant (p-value=0.557, Wilcoxon signed-rank test). When $L$ labels play against other $L$ labels, 41% choose their preferred action in the low payoff game and 38% do so in the high payoff game. The direction is as hypothesized but is insignificant (p-value=0.774, Wilcoxon signed-rank rank test).

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 game. These were included to examine the viability of some alternative mechanisms, such as risk dominance, that may guide subjects' choices in these games and to check whether the basic logic of the model also holds in these games. Our findings for these additional games are discussed in Appendices A.3.3 and A.3.4.

### 5.3 Alternative Models

It may be natural to ask whether our model is equivalent to standard level-$k$ with a mixture of more and less sophisticated players. This is not the case, even under very permissive assumptions about the levels of the subjects. To illustrate, take the BoS game, and assume that the Low label subjects are of lower level than the High label subjects. First, note that this assumption would not suffice to make predictions on behavior. Even so, fixing any level for the subjects, the level-$k$ model would generate the *same* behavior across *all* treatments in our experiment. This is in contrast with the predictions of our model and with the observed behavior. Second, note that if we were to assume that the subjects changed their beliefs over their opponents when facing different labels, then we would have an issue defining level-$k$ behavior. In particular, consider a High player facing another High player. From within the level-$k$ model, it cannot be that a player takes the other player to be of the same level as his own, and so behavior here is not well-specified. Being more permissive with the level-$k$ model, and assuming that when facing the same level they best respond to the level below, we would be back to the issue discussed above – there would be no variation across treatments. Hence, even with a very 'hands-on' approach, a level-$k$ model with a higher and lower cognitive type could not explain the findings. Similarly, QRE would also generate the same distribution across all of the belief treatments (for each payoff specification), unless of course one allows the logit parameter to freely change across treatments, but then this model cannot be used to make predictions across treatments, and it would not be falsifiable. Finally, for some logit parameters multiple QREs exist, and in those cases the QRE-model does not provide a criterion for equilibrium selection, which is one of the central questions of our paper.

### 5.4 Limitations of the Model

Here we briefly discuss potential limitations of the model. Of course, as with any theory, we do not expect it to be a full reflection of reality, but rather one that brings to light potential mechanisms that may be at play.

We first note that in reality, subjects' behavior is influenced by several factors that bring about stochasticity in behavior or other departures from the baseline model. Our maintained assumptions for the experiments accommodate these factors to some extent, but there may well be more types of stochasticity than we account for explicitly.

As an example, for the reverse Strategic Advantage game, observed behavior suggests that some subjects may believe that they are faced with *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). While we could formally introduce such player types within our framework, we have not done so to focus on our main insights.

In this regard, we stress that the main objective of this paper is to bring out a novel theoretical mechanism that may affect individuals' reasoning and ability to attain equilibrium outcomes in standard coordination games, on a purely eductive basis. The objective of the experiment is to test the presence of this mechanism, and to test its key implications for aggregate behavior. That is, it serves as a test of the empirical relevance of a novel theoretical idea, while the main focus of the paper remains conceptual.

In the same spirit, we also note that there may well be agents whose value of reasoning may take very different forms from what we allow for. Such value functions could be incorporated within our framework if we had clear candidates. Moreover, it could well be that the start of the path of reasoning is influenced by how large a payoff is, which we have not accounted for in our model so as to avoid adding a degree of freedom at our disposal. These potentially important factors are left for future research.

## 6 Conclusion

Individuals often face problems in which they must attempt to coordinate with other individuals with whom they rarely interact, with no possibility to communicate and no clear focal points, therefore having only introspective reasoning to resort to. Up to date there has been no mechanism to explain whether or under what conditions coordination might be achieved in these situations. This paper provides such a theory and an experimental test of its predictions.

We show that, even without focal points, coordination is the outcome of a large class of introspective reasoning processes, as long as players view each other as having different cognitive abilities, and that they agree on their relative sophistication. Thus, while it is common to view homogeneity and shared culture as leading to increased coordination (e.g., Kets and Sandroni 2019, Kets et al. 2022, and Kets 2022), in the absence of focal points it is *heterogeneity* that leads to coordination.

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 (cf. Fig. 10) the attribution of the strategic advantage is reversed, in that coordination occurs on the equilibrium that is more favorable to the *more* sophisticated player. First, this shows that our model sheds light on the features of the strategic interaction that determine whether, conditional on being in a heterogeneous matching, it is more beneficial to be perceived as the relatively more or less sophisticated player. Second, this observation clarifies that the predictions of our model are 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 are consistent with the mechanism introduced in this paper.

The discussion above naturally gives rise to questions about individuals' incentives to be perceived as more or less sophisticated, if they had the opportunity to manipulate such perceptions. In many settings, this is not easily done, because perceived sophistication may be due to some characteristic of the group to whom they belong, or because it may be the result of the agent's past behavior in different situations, with different payoff configurations. For instance, a professor would arguably be taken to be more sophisticated than a student, or a more experienced agent than a less experienced agent. That said, for settings in which agents could freely manipulate their perceived sophistication, our theory provides insight into when they would want to appear as more sophisticated and when they would want to appear as less sophisticated. In the BoS game, for instance, they would rather be perceived as less sophisticated, while in the RevSA game they would prefer to be perceived as more. In both games, however, they would prioritize wanting to appear as having a different, rather than similar, degree of sophistication.

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, 2022; Lambrecht, Proto, Rustichini, and Sofianos 2022). This paper shows that such beliefs can play an important role in achieving coordination even in isolated settings. We have focused on simple static games, such as the BoS game, which are the standard archetypes to investigate fundamental questions of coordination, but richer strategic settings (such as repeated games, for instance) may raise further dimensions to players' reasoning, due to the complexity of the game and the possibility of observing and reacting to the opponents' actions, and accounting for their reactions, and so forth. Interestingly, however, despite the important differences between these environments and the games we consider, the qualitative predictions that we obtain from our model are in line with the experimental findings of Proto et al. (2019, 2022) and Lambrecht et al. (2022), on the effects that these beliefs have on the behavior in the repeated BoS.[19] A more systematic extension of our model to dynamic games therefore seems to be a promising avenue for future research.

[19] In these papers, the BoS game is played repeatedly, and hence players can use their moves as a signal of their intelligence (cf. Lambrecht, Proto, Rustichini, and Sofianos 2022). Despite these differences, there there are very interesting similarities in the results. For example, increasing the payoffs' inequalities (tantamount to increasing $r$ in our notation) helps the relatively less intelligent subjects (results 3.5 and 3.6 in Lambrecht et al. 2022).

## References


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

——— (forthcoming): "Coordination and Sophistication [dataset]," *Journal of the European Economic Assocation.*

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.

Guesnerie, R. (2001): *Assessing Rational Expectations: Sunspot multiplicity and economic fluctuations*, MIT Press, Cambridge, MA.

——— (2005): *Assessing Rational Expectations 2: Eductive stability in economics*, MIT Press, Cambridge, MA.

Halevy, Y., J. Hoelzemann, and T. Kneeland (2021): "Magic Mirror on the Wall, Who Is the Smartest One of All?" .

Jagau, S. (2023): "To Catch a Stag: Identifying Payoff-and Risk-Dominance Effects in Coordination Games," *SSRN 4463969*.

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*.

Morris, S. and H. S. Shin (2003): *Global Games: Theory and Applications*, Cambridge University Press, 56–114, Econometric Society Monographs.

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.

——— (2022): "Intelligence, Errors, and Cooperation in Repeated Interactions," *The Review of Economic Studies*, 89, 2723–2767.

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 Additional Regression Tables for BoS Game

| | (1) Choice of X | (2) Choice of X | (3) Choice of X | (4) Choice of X |
|---|---|---|---|---|
| Opponent has H label | 0.204*** | 0.204*** | 0.311*** | 0.311*** |
| | (3.10) | (3.20) | (4.54) | (4.61) |
| Constant | 0.344*** | 0.344*** | 0.368*** | 0.368*** |
| | (7.27) | (7.41) | (7.72) | (7.82) |
| Observations | 209 | 209 | 212 | 212 |

$t$ statistics in parentheses

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

Table 3: BoS Game: Robustness panel (fixed effects) regression results for Hypothesis 1.1, $H$ label players.
Models (1) and (2) give the results for the low payoff versions of the BoS game, with bootstrapped and jackknifed standard errors respectively, and (3) and (4) give the equivalent models for the high payoff versions. Standard errors are clustered at the subject level.

| | (1) Choice of X | (2) Choice of X | (3) Choice of X | (4) Choice of X |
|---|---|---|---|---|
| Opponent has H label | 0.0930 | 0.0930 | 0 | 0 |
| | (0.80) | (0.81) | (0.00) | (0.00) |
| Constant | 0.535*** | 0.535*** | 0.605*** | 0.605*** |
| | (6.86) | (6.95) | (8.37) | (8.01) |
| Observations | 86 | 86 | 86 | 86 |

$t$ statistics in parentheses

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

Table 4: BoS Game: Robustness panel (fixed effects) regression results for Hypothesis 1.2, $L$ label players.
Models (1) and (2) give the results for the low payoff versions of the BoS game, with bootstrapped and jackknifed standard errors respectively, and (3) and (4) give the equivalent models for the high payoff versions. Standard errors are clustered at the subject level.

| | (1) Choice of X | (2) Choice of X |
|---|---|---|
| High Payoff Dummy | 0.126* | 0.0189 |
| | (1.84) | (0.38) |
| Constant | 0.548*** | 0.349*** |
| | (15.74) | (13.89) |
| Observations | 209 | 212 |

*t* statistics in parentheses

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

Table 5: BoS Game: Panel (fixed effects) regression results testing Hypotheses 2.1 and 3.1 ($H$ label players).
Model (1) gives the choice of preferred action $X$ for $H$ label players 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 | (3) Choice of X | (4) Choice of X |
|---|---|---|---|---|
| High Payoff Dummy | 0.126* | 0.126* | 0.0189 | 0.0189 |
| | (1.83) | (1.84) | (0.36) | (0.38) |
| Constant | 0.548*** | 0.548*** | 0.349*** | 0.349*** |
| | (11.21) | (11.10) | (7.35) | (7.50) |
| Observations | 209 | 209 | 212 | 212 |

*t* statistics in parentheses

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

Table 6: BoS Game: Robustness panel (fixed effects) regression results testing Hypotheses 2.1 and 3.1 ($H$ label players).
Models (1) and (2) give the choice of preferred action $X$ for $H$ label players in the BoS game against a $H$ label opponent, with bootstrapped and jackknifed standard errors respectively, while Models (3) and (4) give the equivalent 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 7: BoS Game: Panel (fixed effects) regression results testing Hypotheses 2.2 and 3.2 ($L$ label players).
Model (1) gives the choice of preferred action $X$ for $L$ label players 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.

| | (1) Choice of X | (2) Choice of X | (3) Choice of X | (4) Choice of X |
|---|---|---|---|---|
| High Payoff Dummy | 0.0698 | 0.0698 | -0.0233 | -0.0233 |
| | (0.68) | (0.68) | (-0.25) | (-0.24) |
| Constant | 0.535*** | 0.535*** | 0.628*** | 0.628*** |
| | (6.86) | (6.95) | (8.37) | (8.42) |
| Observations | 86 | 86 | 86 | 86 |

*t* statistics in parentheses

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

Table 8: BoS Game: Robustness panel (fixed effects) regression results testing Hypotheses 2.2 and 3.2 ($L$ label players).
Models (1) and (2) give the choice of preferred action $X$ for $L$ label players in the BoS game against a $L$ label opponent, with bootstrapped and jackknifed standard errors respectively, while Models (3) and (4) give the equivalent results against a $H$ label opponent. Standard errors are clustered at the subject level.

### A.2 Regression Tables for RevSA Game

| | Low payoff games | High payoff games |
|---|---|---|
| | (1) | (2) |
| | Choice of X | Choice of X |
| Opponent has H label | -0.163*** | -0.0286 |
| | (-2.66) | (-0.54) |
| Constant | 0.352*** | 0.256*** |
| | (11.59) | (9.64) |
| Observations | 210 | 211 |

*t* statistics in parentheses

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

Table 9: RevSA Game: Panel (fixed effects) regression results for $H$ label players. Model (1) gives the results for the low payoff versions of the RevSA game and (2) for the high payoff versions. Standard errors are clustered at the subject level.

| | Low payoff games | High payoff games |
|---|---|---|
| | (1) | (2) |
| | Choice of X | Choice of X |
| Opponent has H label | 0.0488 | -0.0476 |
| | (0.39) | (-0.49) |
| Constant | 0.416*** | 0.377*** |
| | (6.42) | (7.71) |
| Observations | 84 | 85 |

*t* statistics in parentheses

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

Table 10: RevSA Game: Panel (fixed effects) regression results for $L$ label players. Model (1) gives the results for the low payoff versions of the RevSA game and (2) for the high payoff versions. Standard errors are clustered at the subject level.

| | (1) | (2) |
|---|---|---|
| | Choice of X | Choice of X |
| High Payoff Dummy | 0.0388 | -0.0943* |
| | (0.78) | (-1.78) |
| Constant | 0.191*** | 0.349*** |
| | (7.65) | (13.18) |
| Observations | 209 | 212 |

*t* statistics in parentheses

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

Table 11: RevSA Game: Panel (fixed effects) regression results testing Payoff Effects ($H$ label players).
Model (1) gives the choice of preferred action $X$ for $H$ label players in the RevSA 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) | (2) |
|---|---|---|
| | Choice of X | Choice of X |
| High Payoff Dummy | -0.0500 | -0.140 |
| | (-0.57) | (-1.42) |
| Constant | 0.423*** | 0.465*** |
| | (9.52) | (9.49) |
| Observations | 83 | 86 |

*t* statistics in parentheses

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

Table 12: RevSA Game: Panel (fixed effects) regression results testing payoff Effects ($L$ label players).
Model (1) gives the choice of preferred action $X$ for $L$ label players in the RevSA 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.3 Additional Results

### A.3.1 Coordination Results

While our experimental hypotheses are for individual behavior, we also analyze 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. Here we do not provide formal hypotheses on coordination comparing homogeneous to heterogeneous treatments for equal payoffs, as they would require stronger assumptions on the comparability of the $L$ and $H$ groups, which we have not imposed.

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 13 and 14).

| $H$ 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 13: Results BoS Game - $H$ label row players, % coordination on the opponents' 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.[20] When $H$ play against other $H$ players, 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 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.

| $L$ 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 14: Results BoS Game - $L$ label row players, % coordination on the player's favorite equilibrium.

Similarly, considering $L$ vs $L$, when we consider $(X, W)$ the percentage of coordination

[20]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 13 and 14 provide the combined percentages, given the interchangeability of the two groups.

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 $H$ versus $H$, 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.

### A.3.2 BoS: Analysis of Noise in Individual Behavior

Under the maintained Assumptions 2, which allow for non-degenerate beliefs and stochasticity, the model does not restrict behavior across treatments at the individual level, but only at the population level. That is the reason why the main experimental hypotheses discussed in the main body of the paper only refer to aggregate behavior.

Nonetheless, it may still be worth asking how individual behavior throughout the experiment compares with the theoretical benchmark of Section 2, where only degenerate beliefs are assumed and all stochasticity is removed. Counting the fraction of departures of individual behavior, compared to this degenerate benchmark, from treatment to treatment, indirectly gives an idea of the level of noise that is present in the data.

In the following, we discuss, for each treatment comparison considered in the hypotheses we tested for aggregate behavior, the conditional probabilities that individual behavior departs from the implication of the benchmark model with degenerate beliefs and no noise. These conditional probabilities are calculated in the following way: For the $H$ label group, conditional on a subject choosing $X$ against a $H$ label opponent in the low payoff version, if (as in the benchmark model) beliefs were degenerate and such that the opponent is strictly more sophisticated, and if (as in the benchmark model) they remain the same across treatments, then they would not switch to playing $Y$ when the payoff increases. If a $H$ subject plays $Y$ against a $L$ label subject, if beliefs were degenerate and such that the opponent is strictly less sophisticated, thus expecting the equilibrium favorable for the $L$ player to occur, and if beliefs were degenerate and do not change (as in the no-noise benchmark) then these subjects should not switch to playing $X$ when payoffs increase. If a $H$ subject already played $Y$ against another $H$ label player with degenerate beliefs that the opponent is strictly more sophisticated, with constant beliefs they should keep doing so when the opponent changes to a $L$ label player. The same holds for the high payoff version of the game. We calculate the conditional probabilities for each and find values of 0.3393, 0.2090, 0.2889, and 0.4063 respectively. The average of these conditional probabilities for the $H$ label subjects is 0.3109.

For the $L$ label subjects, under the equivalent degenerate belief that both $H$ and $L$ opponents are strictly more sophisticated, if someone selected $X$ against a $L$ opponent for the low payoff version of the game, they should not switch to playing $Y$ against a $H$ opponent. Similarly, they should not switch to playing $Y$ when the payoff increases. Further, conditional on having chosen $X$ against a $H$ opponent in the low payoff game, they should not switch to playing $Y$ in the high payoff game against a $H$ opponent. Finally, in the high payoff version of the game, conditional on having chosen $X$ against a $L$ opponent, they should not play $Y$ against a $H$ opponent. The conditional probabilities are 0.4348, 0.3478, 0.3333, and 0.4231 respectively. The average conditional probability of a departure for the $L$ label subjects is 0.3848.

As these conditional probabilities are at the aggregate level, we also consider individual behavior across treatments. Doing so, we note that across all treatments and all comparisons, the fraction of individuals for whom the number of departures from the no-noise benchmark is one or less is around 80%, while around 55% have no departures from that benchmark.

### A.3.3 Stag Hunt

Stag Hunt was included to assess whether risk dominance might provide an alternative explanation for observed behavior.[21] This game, shown in Fig.13, 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 13: The *Stag Hunt* Game, with $r \in \{50, 70\}$.

In the discussion below, we maintain the same identification assumptions as in the text, and obtain the following hypotheses. Note that both the row and the column players have the same preferred action profile. Moreover, as $r$ increases, row players are more likely to stop their reasoning at $X$, with the conjecture being that their opponent will stop at $W$. Hence, as $r$ increases, all subjects are more likely to play $X$, regardless of whether they are driven by their own cognitive bound, or by best responding to the

[21] Note that risk-dominance in the BoS always predicts that players choose the action associated with their most preferred equilibrium. For a thorough analysis of the experimental evidence on Stag Hunt and risk-dominance, see Jagau (2023).

opponents'. This also means that, while we have clear comparative statics as $r$ increases for both homogeneous (same) and heterogeneous (other) treatments, the comparisons from homogeneous to heterogeneous treatments are ambiguous for this game.

Like in the main text, we let $p^l(SH_j)$ denote the fraction of label $l \in \{L, H\}$ subjects playing $X$ in Stag Hunt, where $j \in \{same, other\}$ refers to whether the opponent has the same label as the subject (*same*) or not (*other*).

**Hypothesis 6 (Stag Hunt comparisons)**

1. *Low to high payoff comparisons for homogeneous treatments:* $p^l(SH_{same}) \leq p^l(SH^+_{same})$ *for both labels* $l \in \{L, H\}$.

2. *Low to high payoff comparisons for heterogeneous treatments:* $p^l(SH_{other}) \leq p^l(SH^+_{other})$ *for both labels* $l \in \{L, H\}$.

In general, we find that a large majority of subjects chooses $X$, the action associated with the Pareto dominant equilibrium, with the comparisons being overall consistent with Hypothesis 6. For the High label players, as $r$ is increased, the fraction choosing $X$ goes from 81% to 86% when they play against a High label opponent, and from 75% to 81% against a Low label opponent (see Figure 14). Both changes are in the right direction, but neither is statistically significant (p-values= 0.210 and 0.286 respectively, Wilcoxon signed-rank test). For the Low label players, as $r$ is increased, the fraction choosing $X$ goes from 77% to 88% against a Low label opponent and from 77% to 74% when they play against a High label opponent, and from in the low payoffs version of the game (see Figure 14). The first comparison goes in the expected direction and is statistically weakly significant ((non-exact) p-value= 0.096 and exact p-value= 0.179, Wilcoxon signed-rank test). The second difference goes in the opposite direction to Hypothesis 6, but is not significant (p-value= 1.0, Wilcoxon signed-rank test). In contrast, the latter Regression results for the High and Low labels are given in Tables 15 and 16, respectively.

Note that, in our treatments, $X$ is risk dominant when payoffs are high, but not when they are low. So, it might seem that the increase in the fraction of subjects choosing $X$ when $r$ increases is consistent with risk dominance. However, the results show that a large majority of subjects chooses $X$ even for the low payoff version, and this goes against risk dominance.[22] This result, however, is fully consistent with our model. 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.

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

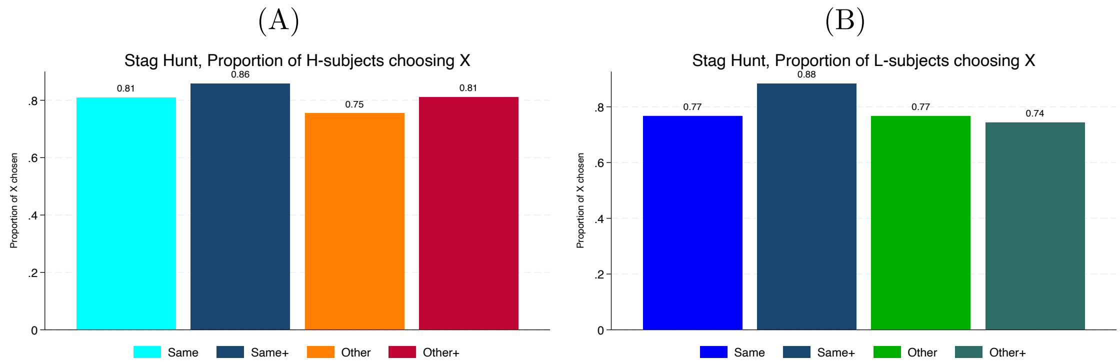



Figure 14: Results for Stag Hunt: Proportion choosing $X$ (their preferred action).
Panel (A) shows the results for the High label Players, and panel (B) for the Low label subjects. *Same* (*Same*+) refers to the low (high) payoff version of the Stag Hunt game played against another player from the same label. *Other* (*Other*+) refers to the low (high) payoff version of Stag Hunt played against another player from the other label.

| | (1)<br>Choice of X | (2)<br>Choice of X |
|---|---|---|
| High Payoff Dummy | 0.0571 | 0.0566 |
| | (1.51) | (1.28) |
| Constant | 0.805*** | 0.755*** |
| | (20.84) | (17.97) |
| Observations | 211 | 212 |

$t$ statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Table 15: Stag Hunt Game: Panel (fixed effects) regression results testing Hypothesis 6 ($H$ label players).
Model (1) gives the choice of preferred action $X$ for $H$ label players in the Stag Hunt 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.

### A.3.4 Asymmetric Matching Pennies

In order to check whether the basic logic of our model also holds in non-coordination games, 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.

Since the incentives are asymmetric between the row and column players, in order to have clear predictions, we must strengthen the assumptions about the distance between a subject's cost of reasoning, $c_i$, and his beliefs about the opponent's, $c_j^i$. Effectively, we need to assume that while the change in $r$ may affect the depth of reasoning of player $i$ (if $r$ affects $i$'s payoffs), or $j$'s depth of reasoning, as *perceived* by $i$ (if $r$ affects $j$'s

| | (1) | (2) |
|---|---|---|
| | Choice of X | Choice of X |
| High Payoff Dummy | 0.116* | -0.0233 |
| | (1.70) | (-0.33) |
| Constant | 0.767*** | 0.767*** |
| | (11.77) | (11.77) |
| Observations | 86 | 86 |

$t$ statistics in parentheses

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

Table 16: Stag Hunt Game: Panel (fixed effects) regression results testing Hypothesis 6 ($L$ label players).
Model (1) gives the choice of preferred action $X$ for $L$ label players in the Stag Hunt 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.

payoffs), it is not enough to revert their relative positions. That is, for each subject, type $(i, c_i, c_j^i)$ is such that, for the two values of $r$ considered in the experiment, it holds that $\mathcal{K}(v_i, c_i) < \mathcal{K}(v_j^i, c_j^i)$ with low $r$ if and only if $\mathcal{K}(v_i, c_i) < \mathcal{K}(v_j^i, c_j^i)$ with high $r$.

With this assumption, the model implies the following experimental hypotheses, where $p^l_{role}(AMP_j)$ denotes the fraction of subjects with label $l \in \{L, H\}$ in each role (row or column, denoted by $role \in \{row, col\}$), that choose $X$ if $role = row$ or $Z$ if $role = col$ against an opponent whose label $j \in \{same, other\}$ is the same as their own or not.

**Hypothesis 7 (Asymmetric matching pennies comparisons)**

1. *For the row players, the following low to high payoff comparisons hold:* $p^l_{row}(AMP_{same}) \leq p^l_{row}(AMP^+_{same})$ *and* $p^l_{row}(AMP_{other}) \leq p^l_{row}(AMP^+_{other})$*, for both labels* $l \in \{L, H\}$.

2. *For the row players, the following same to other label comparisons hold:* $p^H_{row}(AMP^+_{same}) \geq p^H_{row}(AMP^+_{other})$ *and* $p^L_{row}(AMP^+_{same}) \leq p^L_{row}(AMP^+_{other})$

3. *For the column players, the following low to high payoff comparisons hold:* $p^l_{col}(AMP_{same}) \leq p^l_{col}(AMP^+_{same})$ *and* $p^l_{col}(AMP_{other}) \leq p^l_{col}(AMP^+_{other})$*, for both labels* $l \in \{L, H\}$.

4. *For the column players, the following same to other label comparisons hold:* $p^H_{col}(AMP^+_{same}) \leq p^H_{col}(AMP^+_{other})$ *and* $p^L_{col}(AMP^+_{same}) \geq p^L_{col}(AMP^+_{other})$.

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

Figure 15: The *Asymmetric Matching Pennies* (AMP) Game, with $r \in \{41, 160\}$.

Considering first point 1 of the hypothesis, the fraction of High label row subjects that choose $X$ against another High label goes from 57% to 62% as payoffs are increased (p-value= 0.678, Wilcoxon signed-rank test), and from 55% to 60% when they play against

Low label subjects (p-value = 0.629, Wilcoxon signed-rank test). For Low label row subjects instead, the fraction that chooses $X$ goes from 52% to 64% playing against other Low labels (p-value = 0.451, Wilcoxon signed-rank test), and from 60% to 64% when they play against High labels (p-value≈ 1.0 , Wilcoxon signed-rank test).

As for point 2 in the hypothesis, the fraction of High label row subjects that chooses $X$ in the high payoff game goes from 62% to 60% as the opponent changes from the same to the other label (p-value≈ 1.0 , Wilcoxon signed-rank test), while for the Low subjects it stays at 64% in both treatments (p-value = 1.0, Wilcoxon signed-rank test). These results are summarized in Figure 16 and in Tables 17, 18, and 19.

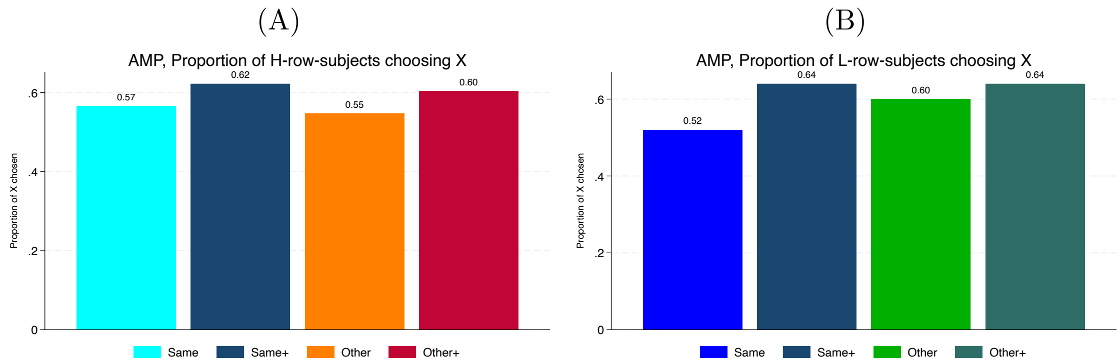



Figure 16: Results AMP: Row Players. Proportion choosing $X$ (their preferred action). Panel (A) shows the results for the High label Players, and panel (B) for the Low label subjects. *Same* (*Same+*) refers to the low (high) payoff version of the AMP game played against another player from the same label. *Other* (*Other+*) refers to the low (high) payoff version of the AMP game played against another player from the other label.

| | (1) | (2) |
|---|---|---|
| | Choice of X | Choice of X |
| High Payoff Dummy | 0.0566 | 0.0566 |
| | (0.62) | (0.72) |
| Constant | 0.566*** | 0.547*** |
| | (12.38) | (13.94) |
| Observations | 106 | 106 |

$t$ statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Table 17: AMP Game: Row Players. Panel (fixed effects) regression results testing Hypothesis 7.1 ($H$ label players).
Model (1) gives the choice of preferred action $X$ for $H$ label players in the AMP 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.

Coming to the column players, for point 3, the fraction of High label subjects that chooses $Z$ against another High label subjects increases from 68% to 74% as payoffs are increased (p-value= 0.648, Wilcoxon signed-rank test), and from 68% to 79% when

| | (1) Choice of X | (2) Choice of X |
|---|---|---|
| High Payoff Dummy | 0.120 | 0.0400 |
| | (1.13) | (0.32) |
| Constant | 0.520*** | 0.600*** |
| | (9.78) | (9.72) |
| Observations | 50 | 50 |

$t$ statistics in parentheses

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

Table 18: AMP Game: Row Players. Panel (fixed effects) regression results testing Hypothesis 7.1 ($L$ label players).
Model (1) gives the choice of preferred action $X$ for $L$ label players in the AMP 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.

| | (H Label) (1) Choice of X | (L label) (2) Choice of X |
|---|---|---|
| Opponent has H label | 0.0189 | 0 |
| | (0.21) | (0.00) |
| Constant | 0.604*** | 0.640*** |
| | (13.16) | (8.96) |
| Observations | 106 | 50 |

$t$ statistics in parentheses

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

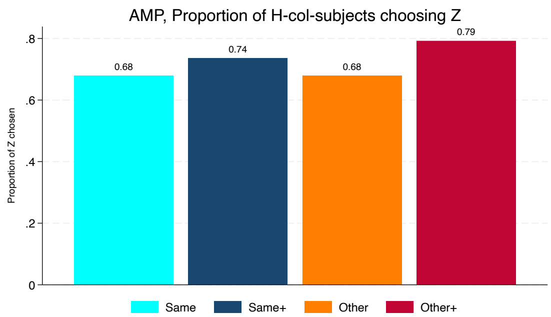
Table 19: AMP Game: Row Players. Panel (fixed effects) regression results testing Hypothesis 7.2 (both labels).
Models (1) and (2) give the label comparison results for the high payoff version of the AMP game, for the H and L labels respectively. Standard errors are clustered at the subject level.

they play against Low labels (p-value= 0.238, Wilcoxon signed-rank test). For Low label subjects instead, the fraction that chooses $Z$ increases from 44% to 72% when playing against other Low labels (p-value= 0.219, Wilcoxon signed-rank test), and from 56% to 67% when they play against High labels (p-value= 0.727, Wilcoxon signed-rank test).

Finally, for point 4, the fraction of High label subjects that chooses $Z$ in the high payoff game goes from 74% to 79% as the opponent changes from the same to the other label (p-value= 0.727, Wilcoxon signed-rank test), while for the Low subjects it decreases from 72% to 67% (p-value≈ 1.0 , Wilcoxon signed-rank test). Figure 17 and Tables 20, 21, and 22 give the results.

Although most of these differences are not statistically significant, all of the comparisons go in the direction predicted by Hypothesis 7. The lack of significance is likely due to the small sample size for this game, as row and column players must be treated separately because of the payoff asymmetry.

(A)




(B)

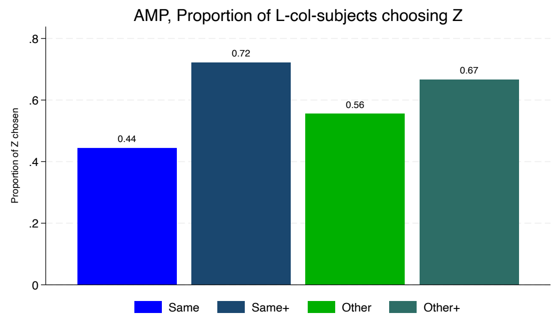



Figure 17: Results AMP: Column Players. Proportion choosing action $Z$.
Panel (A) shows the results for the High label Players, and panel (B) for the Low label subjects. *Same* (*Same+*) refers to the low (high) payoff version of the AMP game played against another player from the same label. *Other* (*Other+*) refers to the low (high) payoff version of the AMP game played against another player from the other label.

| | (1)<br>Choice of Z | (2)<br>Choice of Z |
|---|---|---|
| High Payoff Dummy | 0.0566 | 0.113 |
| | (0.65) | (1.42) |
| Constant | 0.679*** | 0.679*** |
| | (15.55) | (17.05) |
| Observations | 106 | 106 |

$t$ statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Table 20: AMP Game: Column Players. Panel (fixed effects) regression results testing Hypothesis 7.3 ($H$ label players).
Model (1) gives the choice of action $Z$ for $H$ label players in the AMP 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.

### A.3.5 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

| | (1) Choice of Z | (2) Choice of Z |
|---|---|---|
| High Payoff Dummy | 0.278* | 0.111 |
| | (2.02) | (0.69) |
| Constant | 0.444*** | 0.556*** |
| | (6.47) | (6.87) |
| Observations | 36 | 36 |

*t* statistics in parentheses

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

Table 21: AMP Game: Column Players. Panel (fixed effects) regression results testing Hypothesis 7.3 ($L$ label players).
Model (1) gives the choice of action $Z$ for $L$ label players in the AMP 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.

| | (H label) (1) Choice of Z | (L label) (2) Choice of Z |
|---|---|---|
| Opponent has H label | -0.0566 | -0.0556 |
| | (-1.00) | (-0.43) |
| Constant | 0.792*** | 0.722*** |
| | (27.87) | (11.20) |
| Observations | 106 | 36 |

*t* statistics in parentheses

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

Table 22: AMP Game: Column Players. Panel (fixed effects) regression results testing Hypothesis 7.4 (both labels).
Models (1) and (2) give the label comparison results for the high payoff version of the AMP game, for the H and L labels respectively. Standard errors are clustered at the subject level.

(p-value = 0.834 for the Combined Kolmogorov-Smirnov test).

#### A.3.6 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 23 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 23: 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 24 below.

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

Table 24: 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. The order of the games was the same for all subjects. They first saw the main game, the BoS, to prevent that other games influence behavior in the game that is the focus of the analysis. Furthermore, subjects completed the Stag Hunt between the BoS and the RevSA games to ensure that subjects were paying attention to the fact that the BoS and the RevSA were different games. They saw the Asymmetric Matching Pennies game last. While the order of the Stag Hunt and the Asymmetric Matching Pennies games could have been randomized, this order was chosen such that subjects completed the symmetric games first, before exposing them to the asymmetric one. Due to the symmetric nature of the first three games, 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). Figure 18 gives a graphical illustration of the experimental design.

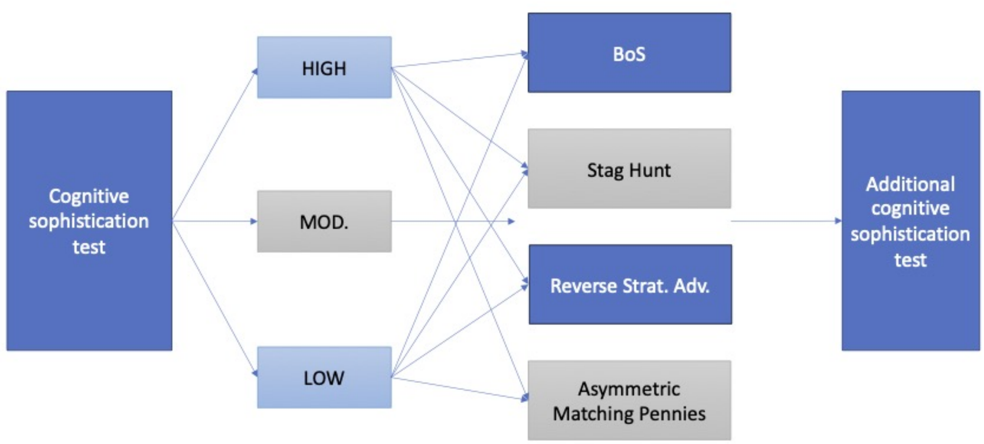



Figure 18: Illustration of the experimental design.

### 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. Text that subjects saw is shown in italics. 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 of the test, the Mastermind game, the Centipede game and the Muddy Faces game, are thus identical.

**Instructions for Our Cognitive Test**

*This test consists of three questions. You must answer all three within the time limit stated.*

**Instructions Mastermind game**

*In this question, you have to guess four numbers in the correct order. Each number is between 1 and 7. No two numbers are the same. You have nine attempts to guess the four numbers. After each attempt, you will be told the number of correct answers in the correct place, and the number of correct numbers in the wrong place.*

*Example: Suppose that the correct number is: 1 4 6 2.*

*If you guess: 3 5 4 6, then you will be told that you have 0 correct answers in the correct place and 2 in the wrong place.*

*If you guess: 3 5 6 4, then you will be told that you have 1 correct answer in the correct place and 1 in the wrong place.*

*If you guess: 3 4 7 2, then you will be told that you have 2 correct answers in the correct place and 0 in the wrong place.*

*If you guess: 1 4 6 2, then you will be told that you have 4 correct answers, and you have reached the objective.*

*Notice that the correct number could not be (for instance) 1 4 4 2, as 4 is repeated twice. You are, however, allowed to guess 1 4 4 2, in any round.*

*You have a total of 90 seconds per round: 30 seconds to introduce the numbers and 60 seconds to view the results.*

**Instructions Centipede game**

*Consider the following game. Two people, Antonio and Beatriz, are moving sequentially. The game starts with 1 euro on the table. There are at most 6 rounds in this game:*

*Round 1) Antonio is given the choice whether to take this 1 euro, or pass, in which case the game has another round. If he takes the euro, the game ends. He gets 1 euro, Beatriz* *gets 0 euros. If Antonio passes, they move to round 2.*

*Round 2) 1 more euro is put on the table. Beatriz now decides whether to take 2 euros, or pass. If she takes the 2 euros, the game ends. She receives 2 euros, and Antonio receives 0 euros. If Beatriz passes, they move to round 3.*

*Round 3) 1 more euro is put on the table. Antonio is asked again: he can either take 3 euros and leave 0 to Beatriz, or pass. If Antonio passes, they move to round 4.*

*Round 4) 1 more euro is put on the table. Beatriz can either take 3 euros and leave 1 euro to Antonio, or pass. If Beatriz passes, they move to round 5.*

*Round 5) 1 more euro is put on the table. Antonio can either take 3 euros and leave 2 to Beatriz, or pass. If Antonio passes, they move to round 6.*

*Round 6) Beatriz can either take 4 euros and leaves 2 to Antonio, or she passes, and they both get 3.*

*Assume Antonio and Beatriz are infinitely sophisticated and rational and they each want to get as much money as possible. What will be the outcome of the game?*
*a) Game stops at Round 1, with payoffs: (Antonio: 1 euro Beatriz: 0 euros)*
*b) Game stops at Round 2, with payoffs: (Antonio: 0 euro Beatriz: 2 euros)*
*c) Game stops at Round 3, with payoffs: (Antonio: 2 euros Beatriz: 1 euro)*
*d) Game stops at Round 4, with payoffs: (Antonio: 1 euro Beatriz: 3 euros)*
*e) Game stops at Round 5, with payoffs: (Antonio: 3 euros Beatriz: 2 euros)*
*f) Game stops at Round 6, with payoffs: (Antonio: 2 euros Beatriz: 4 euros)*
*g) Game stops at Round 6, with payoffs: (Antonio: 3 euros Beatriz: 3 euros)*

*You have 8 minutes in total for this question.*

**Instructions Muddy Faces game**

*There are three people, A, B and C, each with a circle on their forehead. The circle can be white or black. Every person can see the circle on the others' forehead but not the one on their own. In reality, A and C have a white circle and B has a black circle:*

*They are given the following instructions, in this order, and can observe the reaction of the others:*

*If you know that your circle is black, take a step forward. Who will take a step forward?*

*Now, they are informed that at least one of them has a black circle. They are then asked: If you know the color of your circle, take a step forward. Who will take a step forward?*

*They observe the reaction to the previous question (in other words, they see who took a step forward). They are asked: Now that you have seen who stepped forward, if you know the color of your circle, take a step forward. Who will take a step forward? (Include only those new persons who take a step forward, don't include anyone who already took a step forward in the previous questions.)*

**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.*

**Scoring of Raven Test**

The Raven Test was scored by calculating the percentage of correct answers to the twelve matrices that the subjects had to complete.

**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).

### A.5 Glossary

This section contains a glossary of the terminology used to describe the experimental design, hypotheses, and results.

- $BoS$: Battle of the Sexes game.
- $BoS_{same}$: the low payoff version of the Battle of the Sexes game, played against an opponent with the same cognitive sophistication label as the decision-maker.
- $BoS_{other}$: the low payoff version of the Battle of the Sexes game, played against an opponent with a cognitive sophistication label different from the decision-maker's.
- $BoS^{+}_{same}$: the high payoff version of the Battle of the Sexes game, played against an opponent with the same cognitive sophistication label as the decision-maker.
- $BoS^{+}_{other}$: the high payoff version of the Battle of the Sexes game, played against an opponent with a cognitive sophistication label different from the decision-maker's.
- $RevSA$: Reverse Strategic Advantage game.
- $RevSA_{same}$: the low payoff version of the Reverse Strategic Advantage game, played against an opponent with the same cognitive sophistication label as the decision-maker.
- $RevSA_{other}$: the low payoff version of the Reverse Strategic Advantage game, played against an opponent with a cognitive sophistication label different from the decision-maker's.
- $RevSA^{+}_{same}$: the high payoff version of the Reverse Strategic Advantage game, played against an opponent with the same cognitive sophistication label as the decision-maker.
- $RevSA^{+}_{other}$: the high payoff version of the Reverse Strategic Advantage game, played against an opponent with a cognitive sophistication label different from the decision-maker's.
- $AMP$: Asymmetric Matching Pennies game.
- $SH$: Stag Hunt game.