Endogenous Depth of Reasoning*

Larbi Alaoui† Antonio Penta‡
Universitat Pompeu Fabra Dept. of Econ., UW-Madison

March 14, 2014

Abstract

We introduce a model of strategic thinking in games of initial response. Unlike standard models of strategic thinking, in this framework the player's 'depth of reasoning' is endogenously determined, and it can be disentangled from his beliefs over his opponent's cognitive bound. In our approach, individuals act as if they follow a cost-benefit analysis. The depth of reasoning is a function of the player's cognitive abilities and his payoffs. The costs are exogenous and represent the game theoretical sophistication of the player; the benefit instead is related to the game payoffs. Behavior is in turn determined by the individual's depth of reasoning and his beliefs about the reasoning process of the opponent. Thus, in our framework, payoffs not only affect individual choices in the traditional sense, but they also shape the cognitive process itself. Our model delivers testable implications on players' chosen actions as incentives and opponents change. We then test the model's predictions with an experiment. We administer different treatments that vary beliefs over payoffs and opponents, as well as beliefs over opponents' beliefs. The results of this experiment, which are not accounted for by current models of reasoning in games, strongly support our theory. We also show that the predictions of our model are highly consistent quantitatively with well-known unresolved empirical puzzles. Our approach therefore serves as a novel, unifying framework of strategic thinking that allows for predictions across games.

Keywords: cognitive cost – depth of reasoning – higher order beliefs – level-k reasoning – strategic thinking

JEL Codes: C72; C92; D80.

*Alaoui gratefully acknowledges financial support from the Spanish Ministry of Science and Innovation under project ECO2011-25295. We thank the editor and three anonymous referees, as well as Ayala Arad, Ghazala Azmat, Andrew Caplin, Miguel Costa-Gomes, Steven Durlauf, Christian Fons-Rosen, Cristina Fuentes-Albero, Jacob Goeree, Serafin Grundl, Terri Kneeland, Bart Lipmann, Pablo Lopez-Aguilar, Paola Manzini, Marco Mariotti, Lones Smith and especially Vincent Crawford, Nagore Iriberri and Rosemarie Nagel for their thoughtful comments. We also thank audiences at seminars at Barcelona Jocs, Bogazici, Bocconi, BU, Columbia, Harvard-MIT, Maastricht, Northwestern, NYU (CESS), Princeton, Stanford (SITE), Toronto, UCLA, University of Illinois, University of Minnesota, UPF, UW-Madison and several conferences.

†Universitat Pompeu Fabra and Barcelona GSE. E-mail: larbi.alaoui@upf.edu

‡University of Wisconsin-Madison. E-mail: apenta@ssc.wisc.edu

1 Introduction

The relevance of economic incentives to individual behavior has long been recognized, but little is known about the effect of incentives on individuals' reasoning processes in strategic settings. The vast experimental literature on initial responses in games shows that individuals' choices depart systematically from classical equilibrium predictions, and the observed regularities suggest that individuals follow distinct, stepwise reasoning procedures, of which they perform only a few steps.[1] But the step reached, or *depth of reasoning*, may depend on the stakes of the game. Furthermore, beliefs over opponents may also affect choices, and the strategic sophistication according to which individuals play need not coincide with their actual sophistication. Hence, cognition, incentives and beliefs interact in strategic settings. But such an interaction is not accounted for by existing models of strategic thinking. If higher stakes provide boundedly rational agents with the incentives to think harder, then caution should be used in interpreting the distribution of play as measures of cognitive abilities, because such measures would be subject to an endogeneity problem. Moreover, since incentives and beliefs vary in the many situations of economic interest in which individuals face strategic uncertainty, analyzing and modeling how the depth of reasoning changes with the strategic setting is key to expanding the reach of bounded rationality models and improving the predictive power of game theory.

In this paper, we introduce a framework in which players' depth of reasoning is endogenously determined as resulting from a procedure that relates individuals' cognitive abilities to the payoffs of the game. Behavior in turn follows from the individual's depth of reasoning and his beliefs about the reasoning process of the opponent. Thus, in our approach, payoffs not only affect individual choices in the traditional sense, but they also shape the cognitive process itself. We next present an experimental test of our theory. The experimental results reveal that individuals change their behavior in a systematic way as payoffs and opponents change, thereby confirming that incentives and beliefs play an important role in determining the agents' depth of reasoning and level of play. From an empirical viewpoint, our findings confirm that an endogeneity problem is present when players' cognitive bounds are assessed from their behavior in isolated games. Moreover, these findings are consistent with the predictions of our theory and strongly support it. To further demonstrate the reach of our approach, we then add structure to our baseline model and demonstrate that it explains well-known empirical puzzles. In particular, we consider Goeree and Holt's (2001) influential "Ten Little Treasures and Ten Intuitive Contradictions" paper, and show that our model is highly consistent with their results, both qualitatively and quantitatively. This analysis also serves to show how our model can be used to make inferences and sharp predictions that hold across different games.

[1] For a recent survey on the empirical and theoretical literature on strategic thinking see Crawford, Costa-Gomes and Iriberri (2012). Particularly important within this area is the literature on level-k reasoning, first introduced by Nagel (1995) and Stahl and Wilson (1994, 1995). Camerer, Ho and Chong (2004) propose the closely related 'cognitive hierarchy' model, in which level-k types respond to a distribution of lower types, and Goeree and Holt (2004) introduce noise in the reasoning process. Level-k models have been extended to study communication (Crawford, 2003), incomplete information (Crawford and Iriberri, 2007) and other games. For recent theoretical work inspired by these ideas, see Strzalecki (2010), Kets (2012) and Kneeland (2013).

The fundamental feature of our framework is that players act *as if* they weigh the incremental value of additional rounds of reasoning against an incremental cost of learning more about the game from introspection. While the cognitive cost is exogenous, the 'value of reasoning' is connected to the game payoffs. In this model, increasing the stakes of the game provides individuals with stronger incentives to reason, which may induce them to perform more rounds of reasoning. But depth of reasoning need not coincide with the sophistication of the chosen action. When facing opponents that they perceive to be more sophisticated than themselves, subjects play according to their own cognitive bound. But when facing less sophisticated opponents, they play according to less rounds of introspection than their actual cognitive bound. We note that the notion of playing a more sophisticated opponent is natural in this setting, thereby resolving a well-known conceptual difficulty of the level-k approach. Our model further predicts that individuals follow (weakly) more sophisticated behavior when the opponents' incentives to reason are increased, unless their own cognitive bound is binding. Depending on the game, these predictions on the depth of reasoning and behavior translate to stochastic dominance relations in the distributions of actions as incentives, payoffs and beliefs over the opponents are varied.

A cost-benefit approach to modeling the reasoning process has several advantages. In addition to holding intuitive appeal, this approach bridges the study of strategic thinking with standard economic concepts. But as this is an unconventional domain of analysis, the extent to which this approach is useful or empirically relevant is not clear. Investigating its empirical relevance, however, presents one important difficulty: since there is no a priori obvious way of specifying the costs and value of reasoning, it is crucial to isolate the core predictions of the approach, which hold independently of the assumptions on the specific functional forms. For this reason, we first introduce a general framework, with minimal restrictions on the cost and benefit of reasoning.[2] Using this 'detail-free' model, we focus on the interaction between players' incentives to reason, their beliefs and their higher-order beliefs, and show that this model delivers a rich set of testable predictions. The detail-free model therefore provides a coherent yet tractable framework for analyzing the complex interaction between the distinct forces at play, and provides the necessary guidance for the design of an empirical test of the core predictions of the cost-benefit approach.[3]

After introducing the general framework, we present our experimental design. Besides testing the predictions of the detail-free model, the experiment also serves the broader purpose of documenting whether players' steps of reasoning vary systematically as their incentives and beliefs over opponents change.[4] Our design consists of varying the players' beliefs and their

[2]In Alaoui and Penta (2013) we pursue an axiomatic approach to players' reasoning, in which the cost-benefit analysis emerges as a representation. Besides uncovering the fundamental underpinnings of the approach, the axioms enable us to impose structure on the functional forms.

[3]Recent work by Choi (2012) also incorporates a cost-benefit approach in a setting of strategic thinking and learning in networks. We discuss the connections with that paper and other models of strategic thinking in Section 2.4.

[4]With respect to the importance of beliefs on players' actions, a recent experiment by Agranov, Potamites, Schotter and Tergiman (2012) makes the simple but important point that beliefs do change the average number

incentives. We consider two different ways of changing the agents' beliefs over their opponents' cognitive abilities. In both cases, we divide the subjects into two groups whose labels are perceived to be informative about game theoretic sophistication. In the first case, we separate the subjects into two groups by degrees of study. In the second, subjects are required to take a test of our design, and are then separated by their score, which can either be 'high' or 'low'. We then use these labels to vary agents' beliefs over their opponents' cognitive constraints. These changes serve to test the model's predictions that agents play according to a lower depth of reasoning when playing against opponents they take to be less sophisticated. Our theory also allows players to not only take into account the (perceived) sophistication of the opponent, but also the opponent's belief over the player's own sophistication. To account for these higher order beliefs effects, we administer treatments in which subjects classified under a label play against the action that subjects from the other label have played against each other.

To test whether players respond to increased incentives of doing more rounds of introspection in the manner that is predicted by our model, we increase their reward for being 'correct' in their reasoning. We then compare the distributions of the chosen actions across these different treatments. Our results are consistent with the prediction that subjects play according to more rounds of introspection when stakes are increased and when opponents are believed to be more sophisticated. The results are also in line with the predictions over higher order beliefs effects. Our model further allows for the analysis of the experiment's more complex observed patterns. In particular, the observed shifts in distributions as beliefs over opponents are changed are more pronounced when players' incentives to reason are weaker. These findings are indicative of an interaction between changes in incentives and changes in beliefs over opponents that is within the scope of our model.

The experimental results show that individuals change their behavior in a systematic manner that is not endogenized by existing models of strategic reasoning, but that is strongly consistent with our theoretical predictions. These findings therefore establish the importance of accounting for the endogeneity of the depth of reasoning and support the validity of our general approach. We then add structure to the model and use it to derive sharp predictions that hold across games. In particular, we assume a specific functional form for the value of reasoning, which is consistent with the representation derived axiomatically in Alaoui and Penta (2013), and constrain agents' beliefs. We then apply this parametric model to the analysis of the games in the influential paper by Goeree and Holt (2001). Goeree and Holt's findings are intuitive, but difficult to reconcile with standard game theory. Nonetheless, we show that our model does not only fit the qualitative results, it also performs well from a quantitative viewpoint. In particular, using a specification of the model with a single free parameter, we calibrate it to match the data in one of Goeree and Holt's experiments, and derive predictions for the other games. We find that the predictions of the calibrated model are highly consistent with Goeree and Holt's empirical results. Since the 'little treasures' are very different from one

of rounds performed, in a standard beauty contest. Palacios-Huerta and Volij (2009) explore a related point in the dynamic context of the centipede game. They analyze how the number of rounds of backward induction performed by subjects vary when the level of sophistication of the opponent is changed.

another, ranging from Basu's (1994) traveler's dilemma to matching pennies and coordination games, these findings show that our model applies to a broad spectrum of games. From a methodological viewpoint, these results show that, by shifting the focus of the analysis to the comparative statics, our theory allows inferences and predictions that hold across games, and enables us to uncover the deeper mechanisms of strategic thinking.

The paper is structured as follows. Section 2 introduces the general theory and its predictions, Section 3 presents the experimental design and Section 4 analyzes the empirical results. Section 5 adds structure to the model and provides a calibration exercise for Goeree and Holt's (2001) findings. Section 6 concludes.

2 Theory

This section introduces our model. We first describe players' reasoning process, which we take as given, and assume that they follow a stepwise procedure. We then endogenize their depth of reasoning, using a cost-benefit analysis. The number of steps they take is a function of their game theoretical sophistication, which determines the cost of reasoning, and the payoff structure of the game, which determines the benefit. Next, we endogenize the players' *choice*, which depend not only on the number of steps of reasoning they perform, but also on their beliefs about the opponent's cognitive abilities and their higher order beliefs. We model hierarchies of beliefs by means of 'cognitive type spaces'. The cost-benefit analysis and the 'cognitive type space' structure are the two central components which determine players' choices. To illustrate the interaction between payoffs, beliefs and higher order beliefs in determining choice, we present a simplified version of our model and use it to derive the predictions that we will test experimentally. We then describe the general model and close the section by briefly discussing the related theoretical literature.

Throughout this section we will use the following two games as leading examples. The first is a modified version of Arad and Rubinstein's (2012) '11-20' game that we will use in our experiment. We defer the discussion of the properties of this game, and its suitability to our objectives, to Section 3. The second is a standard coordination game with two pure strategy Nash equilibria. It will serve to illustrate the structural exercise of Section 5.

A. The (modified) 11-20 game:

> Two players are asked to simultaneously announce an (integer) number between 11 and 20. Players always receive a number of tokens equal to the number they announce. However, if a player announces a number *exactly one less* than his opponent, then he receives an extra reward of x tokens, where $x \geq 20$. If both players choose the same number, then they both receive an extra 10 tokens. For simplicity, assume that each token corresponds to one unit of payoff.

B. Coordination Game: Consider the game in the following matrix, with $x > 0$:

	L	R
T	$x, 40$	$0, 0$
B	$0, 0$	$40, 40$

2.1 Steps of Reasoning

To keep the notation simple, we focus on two-player games with complete information, $G = (A_i, u_i)_{i=1,2}$, where A_i is the (finite) set of actions of player i and $u_i : A_1 \times A_2 \to \mathbb{R}$ is player i's payoff function. For simplicity, we assume that G is such that the (pure strategy) best response correspondence $BR_i : A_j \to A_i$, defined as

$$BR_i(a_j) = \arg\max_{a_i \in A_i} u_i(a_i, a_j) \text{ for each } a_j \in A_j,$$

is single-valued. We assume that each player's reasoning is represented by a sequence of (possibly mixed) strategy profiles $\left\{\left(a_1^k, a_2^k\right)\right\}_{k\in\mathbb{N}}$ such that $a_i^{k+1} = BR_i(a_j^k)$ for each $k = 0, 1, ...,$ and $i \neq j$.[5] We refer to these sequences as *paths of reasoning*, and to profile $a^0 = \left(a_1^0, a_2^0\right)$ as 'the anchor'. Action a_i^0 is what player i would play by default, without any strategic understanding of the game. As player i performs the first step of reasoning, however, he becomes aware that his opponent could play a_j^0, and thus considers playing $a_i^1 = BR_i(a_j^0)$. Similarly, as player i advances from step $k-1$ to step k, he realizes that his opponent may play a_j^{k-1}, in which case the best response would be $a_i^k = BR_i(a_j^{k-1})$.

We interpret the steps of reasoning as 'rounds of introspection'. In our model, players are not boundedly rational in the sense of failing to compute best responses. Rather, players are limited in their ability to conceive that the opponent may perform the same steps of reasoning.

As an illustration, consider the modified 11-20 game described above. For any player i, action $a_i^0 = 20$ is a natural action for a level-0 player, as it is the number that a non-sophisticated player would report if he ignored all strategic considerations.[6] If player i exerts cognitive effort and performs the first step of the reasoning process, then he realizes that his opponent may play 20, in which case his best response would be 19. If he performs a second step, then he realizes that j may also have performed one step of reasoning, and that j could best respond to 20 by choosing 19. That is, he becomes aware that his opponent may choose 19, in which case his best response would be $a_i^2 = 18$. This reasoning continues until he reaches 11, in which case the best response remains at 11.

This process describes player i's understanding of the game, but it does not necessarily describe his actual play. Player i's action also depends on his beliefs about player j's cognitive abilities. For instance, if player i has performed three steps of reasoning then he understand

[5] In general, the best response need not be unique. In case of multiplicity, we assume that the action is drawn from a uniform distribution over the best responses. We abuse notation and write $a_i^1 = BR_i(a_j^0)$ in both cases.

[6] As we will discuss in Section 3, different specifications of the level-0 (including the uniform distribution) would not affect the analysis. For simplicity, we only consider $a_i^0 = 20$ here.

enough to play 17 *if* he believes that j plays according to two steps of reasoning. But if i thinks that j has performed fewer steps, then i would not play 17.

Two cases are possible. If the anchor a^0 is a Nash Equilibrium, iterating the best replies implies that $a^k = a^0$ for every k. That is, if player i approaches a game with an 'anchor' that specifies a certain equilibrium, such as $a^0 = (T, L)$ in the coordination game above (game B), then further introspection of that initial presumption would not challenge his initial view. If instead a^0 is not a Nash Equilibrium, then $a_i^0 \neq a_i^1$ and the reasoning process generates a path which, depending on the game, may either converge to a Nash equilibrium or enter a loop. For instance, suppose that $a^0 = (T, R)$ in the coordination game example. Then, upon further introspection, player 1 'becomes aware' of the coordination problem, and wonders which of the two actions he should play. As we will discuss, this loop between T and B would only be interrupted if the steps of reasoning become 'too costly', or due to 1's beliefs about the opponent's choice.

2.2 Individual Understanding of the Game.

The model we propose for endogenizing the steps of reasoning taken by players is based on a cost-benefit analysis. Performing additional rounds of reasoning entails incurring a cognitive cost. While these costs reflect a player's cognitive ability, which we view as exogenous, we assume that the benefits of performing an extra step of reasoning depend on the payoff structure of the game. This captures the idea that different games may provide different incentives to think.

We stress that we do not view this cost-benefit analysis as an optimization problem actually solved by the agent, but rather as a modeling device to represent a player's reasoning about the game. We hypothesize that agents' understanding of the game varies systematically with the payoff structure. To the extent that players' understanding of the game exhibits this form of consistency, it can be modeled *as if* the cognitive bound $\hat{k}_i$ results from a cost-benefit analysis. This is formally shown in Alaoui and Penta (2013), where we provide an axiomatic foundation to our approach, and derive the cost-benefit representation from primitive assumptions on the player's reasoning process, explicitly modeled as a Turing machine.

2.2.1 Cognitive Costs and Value of Introspection.

Formally, we assume that the value of doing extra steps of reasoning only depends on the payoff structure of the game. Fixing the game payoffs, we define function $v_i : \mathbb{N} \to \mathbb{R}_+$, where $v_i(k)$ represents i's value of doing the k-th round of reasoning, given the previous $k-1$ rounds. The cognitive ability of agent i is represented by a cost function $c_i : \mathbb{N} \to \mathbb{R}_+$, where $c_i(k)$ denotes i's incremental cost of performing the k-th round of reasoning. The following assumptions on c_i and v_i are maintained throughout.

Condition 1 *Maintained assumptions on the Cost and Value of Reasoning:*

1. ***Cost of Reasoning:*** $c_i(0) = 0$ *and* $c_i(k) \geq 0$ *for every* $k \in \mathbb{N}$.

2. ***Value of Reasoning:*** *The value of reasoning only depends on the payoffs of the game, and* $v_i = v_j$ *if the game is symmetric. Furthermore,* $v_i(k) \geq 0$ *for every* $k \in \mathbb{N}$.

It is useful to introduce the following mapping, which identifies the intersection between the value of reasoning and the cost function: Let $\mathcal{K} : \mathbb{R}_+^{\mathbb{N}} \times \mathbb{R}_+^{\mathbb{N}} \rightarrow \mathbb{N}$ be such that, for any $(c, v) \in \mathbb{R}_+^{\mathbb{N}} \times \mathbb{R}_+^{\mathbb{N}}$,

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

with the understanding that $\mathcal{K}(c, v) = \infty$ if the set in equation (1) is empty. Player i's *cognitive bound*, which represents his understanding of the game, is then determined by the value that this function takes at (c_i, v_i):

Definition 1 *Given cost and value functions* (c_i, v_i)*, the* cognitive bound *of player* i *is defined as:*

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

Player i therefore stops the iterative process when the value of performing an additional round of introspection exceeds the cost. The point at which this occurs identifies his cognitive bound $\hat{k}_i$. Note that a player does not compare the benefits and costs at higher k's. That is, he does not consider stages of reasoning higher than his current one. A player who has performed k rounds of introspection is only aware of the portion uncovered by the k steps, and performs a 'one-step ahead' comparison of the incremental cost and value of reasoning.

This process may appear to translate to a standard optimization problem, in which an agent's marginal cost-marginal benefit analysis can be interpreted as first order conditions of a 'total' value and cost tradeoff. If the functions c_i and v_i cross at most once, then the two procedures are indeed the same. But if they cross more than once, then, with a standard optimization problem, the first crossing would not necessarily be the optimal one. Thus, player i does not 'optimize' globally by taking into account the entire curves; rather, he optimizes locally. The 'myopic' (one-step-ahead) behavior that we assume captures the idea, inherent to the very notion of bounded rationality, that the agents do not know (or are not aware of) what they have not yet thought about. Formalizing this notion is often perceived to be a fundamental difficulty in developing a theory of bounded rationality; in this model it emerges naturally.

Whereas the cost function represents the cognitive ability of the player for that game, function v_i represents the perceived value of performing each extra round of reasoning in the game. We take this value to be purely instrumental to informing the player's choice of action, and we assume that the cost-benefit analysis conducted by the agent, and the ensuing bound $\hat{k}_i$, are independent of the opponent's identity, for a given game and a given reasoning procedure.

The v_i function thus depends only on the payoff structure of the game. In this respect, the bound $\hat{k}_i$ can be seen to be determined separately from the player's beliefs about the opponent's sophistication, although both factors affect his behavior. This captures the notion that the function v_i represents the agent's understanding of the game itself, and that his opponent cannot affect this understanding. Consider, for instance, an agent who decides on whether to perform one step of reasoning. He is unaware of what he does not yet understand, but he knows that what he might learn after one step depends only on the game structure. In the modified 11-20 game, when considering the first step, it does not matter for his understanding of the game who his opponent is. Note that this statement is in fact 'correct': what he will actually learn from performing the step is that 19 is a best-response to 20, irrespective of the sophistication of his opponent. Whether this will be relevant for the agent's choice depends on his beliefs about the opponent, which we introduce in the next subsection. This observation also justifies the assumption that $v_i(k) \geq 0$ (Condition 1): net of its cost, a deeper understanding of the game is never detrimental to the agent, who can at worst ignore the extra insight of each further step of reasoning.[7] By performing the second step of reasoning, i realizes that j could play 19, in which case i would best respond by playing 18. But i is not forced to play 18; he can play according to less rounds if he believes that j has not performed the first step of reasoning.

Relative Sophistication. Players can be heterogeneous in their cognitive abilities. This is captured by different cost functions:

Definition 2 *Consider two cost functions, c' and c''. We say that cost function c' is 'more sophisticated' than c'', if $c'(k) \leq c''(k)$ for every k.*

We do not define sophistication directly in terms of the cognitive bounds $\hat{k}_i$ and $\hat{k}_j$ because these bounds are a consequence of the cost-benefit analysis, which is determined endogenously by the game. When payoffs are different for the two players, it may be that $\hat{k}_i < \hat{k}_j$ even if i is more sophisticated than j, in the sense of Definition 2. For instance, this may hold if player i has lower incentives than player j.

We separate the space of cost functions as follows. For any $c_i \in \mathbb{R}_+^{\mathbb{N}}$, let

$$C^+(c_i) = \left\{ c' \in \mathbb{R}_+^{\mathbb{N}} : c_i(k) \geq c'(k) \text{ for every } k \right\} \text{ and}$$
$$C^-(c_i) = \left\{ c' \in \mathbb{R}_+^{\mathbb{N}} : c_i(k) \leq c'(k) \text{ for every } k \right\}.$$

Thus, based on Definition 2, $C^+(c_i)$ and $C^-(c_i)$ are comprised of the cost functions that are respectively 'more' and 'less' sophisticated than c_i.

[7]Note that j does not observe the rounds of reasoning as i performs them, and so additional reasoning cannot have negative value from becoming common knowledge. We stress that v_i does not represent the 'actual' gain of performing extra steps of reasoning, which is unknown to player i and which depends on the opponent's behavior. In this 'initial response' setting, the value of reasoning need not coincide with the realized gain in payoffs.

Cognitive Equivalence. In what follows, we make comparisons of behavior and depth of reasoning across strategic settings. In particular, we analyze how changing payoffs affect players' cognitive bound through changes in the value of reasoning. Phrased differently, we conduct a comparative statics exercise on $\mathcal{K}(c_i, v_i)$ as v_i is changed. But for this comparative statics exercise to be meaningful, it is important to shift v_i without shifting the cost function c_i. For instance, we would not compare the 11-20 game with low stakes to the normal form of chess with high stakes, and conclude that the higher incentives imply that the depth of reasoning must be higher in chess. This logic flawed because the cost of reasoning is arguably much higher in chess, and hence both the cost and the value of reasoning vary in the same direction, so that the overall effect on the cognitive bound is ambiguous.

We avoid this issue by comparing games that are sufficiently similar from a cognitive viewpoint that they entail the same cognitive cost. For instance, the 11-20 game with $x = 20$ or $x = 80$ (the extra payoff for being exactly one below the opponent), have essentially the same structure, and so are equally difficult to understand, even though they may provide different incentives to reason. The following notion of *cognitive equivalence* formalizes the idea:[8]

Definition 3 *Games* $G = (A_i, u_i)_{i=1,2}$ *and* $G' = (A'_i, u'_i)_{i=1,2}$ *are* cognitively equivalent *if, for each* $i \in \{1,2\}$, $A_i = A'_i$ *and the paths of reasoning associated with each game are identical, i.e.* $\{a^k\}_{k\in\mathbb{N}} = \{a'^k\}$.

Consistent with the axiomatic foundation in Alaoui and Penta (2013), we assume the following:

Condition 2 *Let* $G = (A_i, u_i)_{i=1,2}$ *and* $G' = (A_i, u'_i)_{i=1,2}$ *be two cognitively equivalent games, with (common) path of reasoning* $\left\{a^k\right\}_{k\in\mathbb{N}}$ *and cost and value of reasoning* (c_i, v_i) *and* (c'_i, v'_i)*, respectively. Then:*

1. *For each* $k \in \mathbb{N}$: $c_i(k) = c'_i(k)$.
2. *For each* $k \in \mathbb{N}$: $v_i(k) = v'_i(k)$ *if* $(u'_i(a_i, a_{-i}) - u'_i(a_i^{k-1}, a_{-i})) = (u_i(a_i, a_{-i}) - u_i(a_i^{k-1}, a_{-i}))$ *for all* $(a_i, a_{-i}) \in A$.
3. *For each* $k \in \mathbb{N}$: $v_i(k) \geq v'_i(k)$ *if* $(u'_i(a_i, a_{-i}) - u'_i(a_i^{k-1}, a_{-i})) \geq (u_i(a_i, a_{-i}) - u_i(a_i^{k-1}, a_{-i}))$ *for all* $(a_i, a_{-i}) \in A$.

Condition 2.1 states that games that are cognitively equivalent are associated with the same cost of reasoning, effectively grouping together games for which reasoning is equally difficult. This way, differences between cognitively equivalent games will only determine differences in the value of reasoning, if any, thereby allowing meaningful comparative statics. So, for instance, the cost of reasoning of a particular player would be the same in the 11-20 game, for different values of $x \geq 20$, and it would be the same in the coordination game for any $x > 0$. The two games, however, are not cognitively equivalent to each other, and therefore may be associated with different costs of reasoning.

[8] Alaoui and Penta (2013) derive this notion from primitive elements of the reasoning process.

Condition 2.2 states that the incentives to reason vary only if payoff differences vary. For instance, if G and G' have the same payoffs up to a constant, then they entail the same incentives to reason. The value of reasoning therefore depends on the payoff variations, not on the levels, and it is higher the more the player's payoff varies with his own or with the opponent's action (Condition 2.3). In our two leading examples, this implies that the value of reasoning at every step is (weakly) increasing in x.

2.2.2 Discussion of the Maintained Assumptions of the 'Detail-Free' Model

Conditions 1 and 2 entail minimal restrictions on the cost and value of reasoning functions. In particular, these conditions contain virtually no assumptions about their shape (their monotonicity, convexity, etc.). Maintaining this level of generality allows us to focus on the essential features of our approach and to capture different kinds of plausible cost functions. For instance, in the modified 11-20 game, a player who understands the inductive structure of the problem would have a non-monotonic cost c_i. His first rounds of reasoning would be cognitively costly but those following the understanding of the recursive structure would not be, as described in Example 1. We do not assume that the cost function has this shape for the experiment, but we allow it.

Clearly, stronger assumptions would enable sharper predictions. Deriving falsifiable predictions that do not depend on parametric assumptions, however, is key to isolate the conceptual and empirical relevance of the cost-benefit approach in this novel domain. We will show that, once agents' beliefs are modeled, this minimal set of assumptions will enable a rich set of testable (i.e. falsifiable) predictions. Stronger restrictions on the functional forms will be imposed for the calibration exercise of Section 5.

Notice that, irrespective of the shape of these functions, the cognitive bound $\hat{k}_i$ is monotonic in players' sophistication and in the incentives to reason: $\hat{k}_i$ (weakly) decreases as the cognitive costs increase, and it (weakly) increases as the value of reasoning increases.

Proposition 1 (Depth of Reasoning) *Under the maintained assumptions of Condition 1: (i) For any c_i, $v_i'(k) \geq v_i(k)$ for all k implies $\mathcal{K}(c_i, v_i') \geq \mathcal{K}(c_i, v_i')$; (ii) For any v_i, $c_i'(k) \geq c_i(k)$ for all k implies $\mathcal{K}(c_i, v_i) \geq \mathcal{K}(c_i', v_i)$.*

This result is immediate, as shown in the following example of players' reasoning process.

Example 1 *In Figure 1, the cost function c_i is non-monotonic and the value of reasoning v_i is constant, but these shapes are chosen for illustrative purposes only. Player i's cognitive bound, $\hat{k}_i$, is determined by the first intersection of c_i and v_i, as in Definition 1. In the graph on the left, $\hat{k}_i = 2$, meaning that player i has 'become aware' of one round of reasoning of the opponent. The grey area represents player i's 'unawareness region' about the opponent's steps of reasoning of level higher than 1. As the value v_i increases, $\hat{k}_i$ remains constant at first, but then increases to $\hat{k}_i' = 3$ when level v_i' is reached. Correspondingly, the grey area of unawareness shifts to the right, uncovering one more round of reasoning of the opponent. If v_i is further*

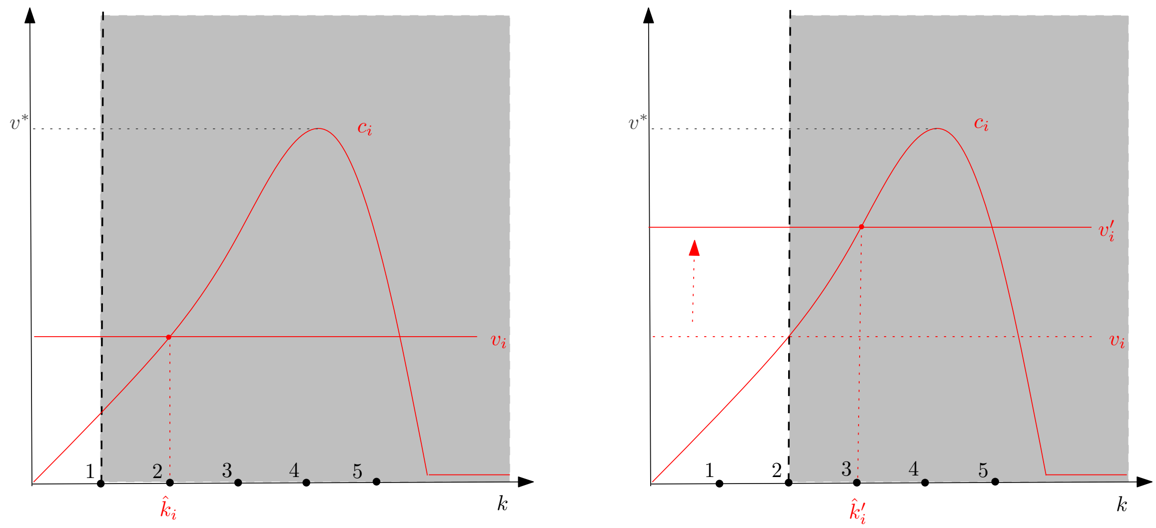


Figure 1: As v_i increases (from the left figure to the right figure), $\hat{k}_i$ (weakly) increases. In both figures, the grey area represents the 'unawareness region' of player i.

increased, i's cognitive bound $\hat{k}_i$ *eventually increases to* 4 *once* v^* *is reached, after which* $\hat{k}_i$ *jumps to* ∞.

The non-monotonic cost function c_i *thus captures the situation of a player who suddenly understands the game after having performed a few rounds of reasoning. At the other extreme, if* c_i *were vertical after any* k*, then there would be an absolute bound, which would not be affected by an increase in* v_i.

2.3 From Reasoning to Choice

As discussed, the cognitive bound of a player does not necessarily determine his behavior. Given a player's understanding of the game, his action also depends on his beliefs about the opponent. To derive the behavioral predictions of the model, it is therefore necessary to complete the model with a specification of agents' beliefs. But if the agent's choices depend on his beliefs about the opponent, then they may also depend on his beliefs about his opponent's beliefs about *him*, and so forth. That is, disentangling depth of reasoning from beliefs about opponents requires accounting for higher order reasoning as well.

Reconciling higher order uncertainty with endogenous depth of reasoning raises important modeling challenges. To fully address the conceptual issues, we provide next a general formulation which serves to analyze potentially complex hierarchies of beliefs. Simple hierarchies of beliefs, however, suffice to understand the theoretical underpinnings of the experiment, in which we vary incentives, beliefs as well as higher order beliefs. For this reason, we introduce the simplified setting in Section 2.3.1, and defer the discussion of the general model to Section 2.3.2.

General Formulation. We model hierarchies of beliefs by means of 'cognitive type spaces'. We first introduce the formal definitions, and discuss the intuition and connection between types and hierarchies in the next subsections, which also contain illustrative examples.

Definition 4 *A 'cognitive type space' (CTS) is a tuple* $\left(T_i, (c_{t_i}, \beta_{t_i})_{t_i \in T_i}\right)_{i=1,2}$ *s.t.* T_i *is a finite set of types of player* i*, and for each type* t_i*,* $c_{t_i} : \mathbb{N} \to \mathbb{R}_+$ *and* $\beta_{t_i} \in \Delta(T_{-i})$ *denote type* t_i*'s cost of reasoning and beliefs about the opponents' types, respectively.*

Given the payoffs of the game and the associated value of reasoning v_i, each type $t_i \in T_i$ in a cognitive type space induces a depth of reasoning $\hat{k}_{t_i} = \mathcal{K}(c_{t_i}, v_i)$. The following definition pins down types' choices, given an anchor a^0 of the reasoning process:

Definition 5 *Fix a cognitive type space and an anchor* a^0*. For each* i*, define function* $\alpha_i^0 : T_i \to A_i$ *such that* $\alpha_i^0(t_i) = a_i^0$ *for each* $t_i \in T_i$*. Recursively, for each* $i = 1, 2$*, for each* $t_i \in T_i$ *and for each* $k \in \mathbb{N}$*, define:*

$$\alpha_i^k(t_i) = \begin{cases} BR_i\left(\sum_{t_{-i} \in T_{-i}} \beta_{t_i}(t_{-i}) \cdot \alpha_{-i}^{k-1}(t_{-i})\right) & \text{if } k \leq \mathcal{K}(c_{t_i}, v_i) \\ \alpha_i^{k-1}(t_i) & \text{otherwise.} \end{cases} \tag{3}$$

The choice of type t_i*, given anchor* a^0*, is* $\hat{a}_i(t_i) := \alpha_i^{\mathcal{K}(c_{t_i}, v_i)}(t_i)$*.*

For the simple CTSs considered in the next section, Definition 5 has a simpler equivalent formulation. The workings of the general recursion (3) are explained in Section 2.3.2.

2.3.1 Simplified Model: Degenerate Beliefs and Second-Order Types

Players' beliefs about their opponents' sophistication need not be correct, as we do not seek for an equilibrium concept and correctness of beliefs is not guaranteed by introspection alone. Therefore, the natural units of analysis are individuals, and particularly their reasoning process and their beliefs, as represented by types in a cognitive type space. Types should thus be regarded in isolation, player by player and type by type.[9]

The general model allows for complex higher order beliefs. In this section we focus on a simple class of types, *second-order types with degenerate beliefs*, which are pinned down by three objects: the cost function c_i, the beliefs about the opponent's, c_j^i, and the beliefs about the opponent's beliefs, c_i^{ij}. These simple types suffice to illustrate the effects that beliefs and higher order beliefs have on behavior, and for the model's predictions for the experiment in Section 3. We first discuss the way in which these simple types, characterized by the triple (c_i, c_j^i, c_i^{ij}), can be formally represented within our general formulation.

Formal discussion of second-order types. Formally, a 'second-order type' for player i is any hierarchy of beliefs that can be represented by a model with the following simple structure:

[9] This approach, also known as the *interim approach*, is the standard one to study non-equilibrium concepts with incomplete information (see, e.g., Weinstein and Yildiz (2007, 2013) or Penta (2012, 2013)).

player i can be one of two types, $T_i = \{c_i, c_i'\}$, whereas player j has only one type $T_j = \{c_j\}$. Type c_j attaches probability q to type c_i, and $(1-q)$ to type c_i'. (Player i's types attach probability one to the only type of player j, c_j.) In a model with degenerate beliefs, q can take two values: 0 or 1.

Each type in a cognitive type space provides a full representation of a player's hierarchy of beliefs. Type c_i, for instance, represents a situation in which player i's cost of reasoning is c_i, and his beliefs about j, which we denote by c_j^i, are $c_j^i = c_j$. If $q = 1$, player i's second order beliefs, denoted by c_i^{ij}, are degenerate and such that $c_i^{ij} = c_i$. In this case, type c_i is a *common belief type*: it represents the situation in which player i thinks that both players believe that they both believe, ..., that the costs of reasoning are, respectively, c_i and c_j. If $c_j^i \in C^-(c_i)$, for instance, player i believes that his opponent is less sophisticated, and that this is common belief. If instead $q = 0$, then c_i's second order beliefs are such that $c_i^{ij} = c_i' \neq c_i$. That is, this type believes that player j believes that i's cost function is different from what it actually is. With $q = 0$ therefore c_i does not represent a 'common belief' situation, and captures player i's concern about j's beliefs being incorrect. In this sense, it represents a situation of higher order uncertainty. For instance, if $c_j^i \in C^-(c_i)$ but $c_i^{ij} \in C^-(c_j^i)$, then i believes that the opponent is less sophisticated, but that he thinks that i is even less sophisticated.

Since any second-order type with degenerate beliefs is characterized by a triple (c_i, c_j^i, c_i^{ij}), in the rest of this section we write types directly as $t_i = (c_i, c_j^i, c_i^{ij})$. We focus on the properties of the c_i, c_j^i and c_i^{ij} functions, which, together with the value functions v_i and v_j, determine the players' depth of reasoning and their *beliefs* over opponents' depth of reasoning and beliefs, as well as players' 'behavioral level' and their beliefs over the opponents. Formally, recall that $\hat{k}_i$ is i's cognitive bound, which is at the intersection of his cost function c_i and his value function v_i (Definition 1). We also define $\hat{k}_j^i$ and $\hat{k}_i^{ij}$ to be i's beliefs over j's cognitive bound and his beliefs over j's beliefs over his (i)'s cognitive bound, respectively. Similarly, define k_i, k_j^i and k_i^{ij} to be i's level of play (or behavioral level), his beliefs over j's level of play, and his beliefs over j's beliefs over i's behavioral level.

Incentives, Beliefs and Behavior. Let (v_i, v_j) denote the value of reasoning for players i and j in a specific game, and let player i's type be $t_i = (c_i, c_j^i, c_i^{ij})$. Player i's beliefs about his opponent's cognitive bound, $\hat{k}_j^i$, is at the intersection of cost function c_j^i and value function v_j *if* he is aware that this occurs. But the maximum bound that he can conceive of for his opponent is constrained to be within the limit of i's own understanding, which is the 'region of awareness' up to $\hat{k}_i - 1$. Hence, player i's belief about $j's$ bound is:

$$\hat{k}_j^i = \min\left\{\hat{k}_i - 1, \mathcal{K}(c_j^i, v_j)\right\}. \tag{4}$$

Suppose, for now, that i is a *common belief type*, i.e. $c_i^{ij} = c_i$. Then, applying Definition 5, i effectively believes that the behavioral level of player j coincides with his cognitive bound, as perceived by i (that is, $k_j^i = \hat{k}_j^i$). Therefore, player i's own behavioral level k_i, which

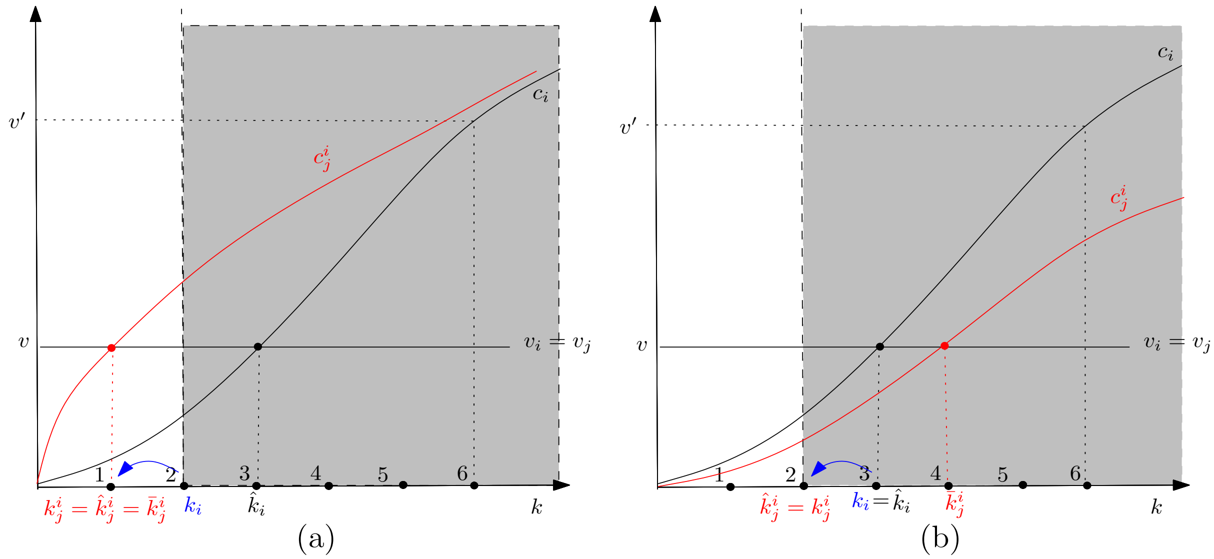


Figure 2: Reasoning about the opponents: on the left, $c_j^i \in C^-(c_i)$; on the right, $c_j^i \in C^+(c_i)$. The grey area represents the 'unawareness region' of player i. The intersection of c_j^i and v_j is denoted $\bar{k}_j^i$.

best-responds to k_j^i, is

$$k_i = k_j^i + 1 = \hat{k}_j^i + 1, \tag{5}$$

with associated action $a_i^{k_i}$. Whether i's own cognitive bound $\hat{k}_i$ constrains his behavioral bound k_i depends on whether i believes that he has performed more or less rounds of introspection than j, as the next example illustrates.

Example 2 *In Figure 2.a, player i, with cost function c_i, perceives his opponent to be less sophisticated. Since the intersection between c_j^i and v_j falls in the region already uncovered by i's cognitive bound, i's belief about j's cognitive bound is at that point, i.e. $\hat{k}_j^i = 1$. This also represents i's belief about j's behavior, k_j^i, hence player i best responds by playing the action associated with level $k_i = k_j^i + 1 = 2$. The cognitive bound $\hat{k}_i$ is not binding, since $k_i < \hat{k}_i$.*

Figure 2.b instead represents the same player reasoning about an opponent that he perceives to be more sophisticated. In this case, the intersection between c_j^i and v_j (denoted $\bar{k}_j^i$ in the graph) falls in the 'unawareness region' of player i. Hence, his perceived cognitive bound for player j is not $\bar{k}_j^i$ but $\hat{k}_j^i = \hat{k}_i - 1$. Player i best responds by playing according to level $k_i = \hat{k}_j^i + 1$, that is, according to his own cognitive bound $\hat{k}_i$.

The next proposition follows from the logic of the example. The common belief type assumption, which we maintain for simplicity, can be weakened for the results.

Proposition 2 (Beliefs and Incentives) *Let* $t_i = (c_i, c_j^i, c_i)$ *be a* common belief type*, and let* $v_i = v_j$*.*

1. *If* $c_j^i \in C^+(c_i)$*, then* $k_i = \hat{k}_i$*. If only* v_i *increases then* $k_i = \hat{k}_i$ *(weakly) increases; if only* v_j *increases then* $k_i = \hat{k}_i$ *does not change; if both* $v_i = v_j$ *increase (preserving the symmetry) then* $k_i = \hat{k}_i$ *(weakly) increases.*

2. *If* $c_j^i \in C^-(c_i)$*, then* $k_i \leq \hat{k}_i$*. If only* v_i *increases then* $\hat{k}_i$ *weakly increases but* k_i *does not change; if only* v_j *increases then* k_i *(weakly) increases and* $\hat{k}_i$ *remains the same; if both* $v_i = v_j$ *increase (preserving the symmetry) then* $k_i = \hat{k}_i$ *(weakly) increases.*

In words, in a game with symmetric value of reasoning, the cognitive bound $\hat{k}_i$ is always binding for a player who believes that he is playing against a more sophisticated opponent. If instead he perceives his opponent to be less sophisticated, then his behavioral k_i is lower than his cognitive bound $\hat{k}_i$. This further implies that player i plays according to a (weakly) deeper k_i when facing a more sophisticated opponent than a less sophisticated opponent. Moreover, changing the value of reasoning of the opponent, while holding his own constant, changes the player's behavior only if he believes that his opponent is less sophisticated. Lastly, if all players' value of reasoning increase then i's cognitive bound $\hat{k}_i$ and his behavioral k_i both increase.[10]

Higher Order Effects and Behavior. Equation (5) derives player i's behavioral level under the assumption that i is a common belief type (i.e., $c_i^{ij} = c_i$). Then, Proposition 2 describes the effects of changing incentives and beliefs about the opponents, holding second-order beliefs fixed. In general, however, the choice of a player depends on his beliefs about the opponent's beliefs about *him*. For instance, if player i is playing an opponent that he regards as less sophisticated, his action may depend on whether or not he believes that the opponent agrees that i is the relatively more sophisticated player. We therefore consider general second-order types, $t_i = (c_i, c_j^i, c_i^{ij})$, without assuming that $c_i^{ij} = c_i$, and then study the effects of changing i's second-order beliefs, c_i^{ij}, while holding c_j^i fixed (Proposition 3).

Formally, while i's beliefs over j's cognitive bound, $\hat{k}_j^i$, do not depend on c_i^{ij}, his beliefs over j's level of play k_j^i do. In particular, k_j^i may be less than $\hat{k}_j^i$ if i believes that j underestimates i's sophistication. In other words, player i puts himself in j's 'shoes', to the extent that he can, and perceives j's beliefs over his own cognitive bound, $\hat{k}_i^{ij}$, to be:

$$\hat{k}_i^{ij} = \min\left\{\mathcal{K}(c_i^{ij}, v_i), \hat{k}_j^i - 1\right\}. \tag{6}$$

Player i then expects j to play according to level $\hat{k}_i^{ij} + 1$, provided that he is capable of conceiving of such a level, which is the case if $\hat{k}_i^{ij} + 1 \leq \hat{k}_i - 1$. Otherwise, he is limited by his

[10] The full separation between players' understanding of the game and their reasoning about the opponent is an important feature of our model. Alternatively, it may appear plausible that the player stops reasoning if he believes that his opponent has already reached his bound, because the extra steps of reasoning would not affect the player's own choice. This alternative formulation can be easily accommodated in our model, and only entails a reinterpretation of some of the variables.

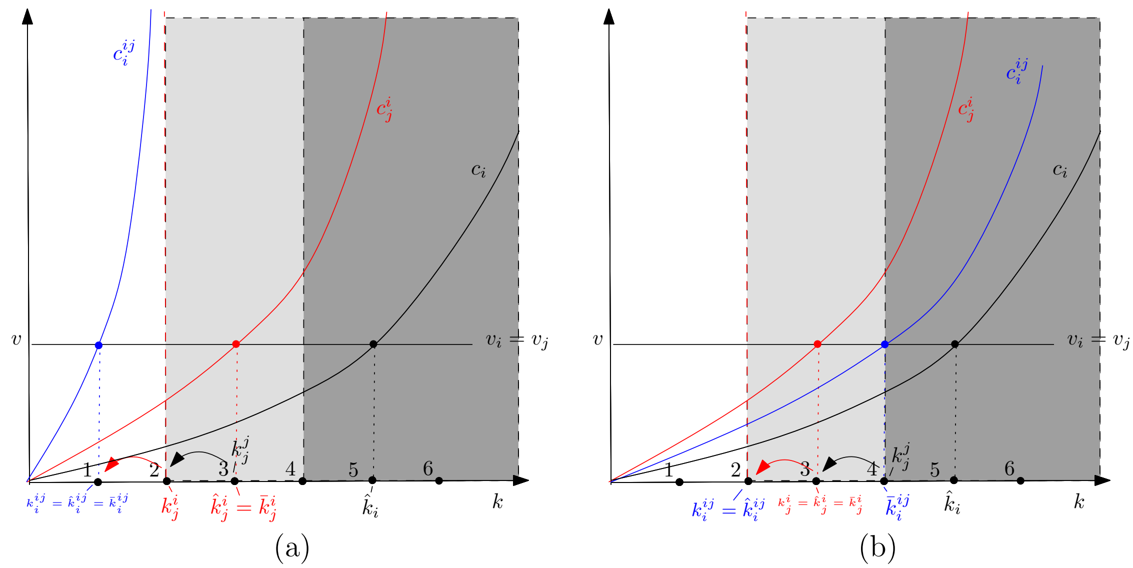


Figure 3: Higher Order Reasoning: $c_j^i \in C^-(c_i)$, with $c_i^{ij} \in C^-(c_j^i)$ on the left, and with $c_i^{ij} \in C^+(c_j^i)$ on the right. The dark grey area represents the 'unawareness region' of player i, whose cognitive bound is $\hat{k}_i = 5$. The light grey area represents the unawareness region of j, as perceived by i. The intersection of c_j^i and $v_j$is denoted $\bar{k}_j^i$, and the intersection of $c_i^{ij}$and v_j is denoted $\bar{k}_i^{ij}$.

own cognitive bound. Hence, for a general second-order type, i's perception of j's behavioral bound is:

$$k_j^i = \min\left\{\hat{k}_i^{ij} + 1, \hat{k}_i - 1\right\}. \tag{7}$$

Player i then best responds by playing action $a_i^{k_i}$, where $k_i = k_j^i + 1$.[11]

Example 3 *Figure 3 represents a player with cost function c_i reasoning about an opponent that he regards as less sophisticated. In Figure 3.a, player i believes that j thinks that i is even less sophisticated (that is, $c_i^{ij} \in C^-(c_j^i)$). Rather than best respond to his perception of j's cognitive bound, $\hat{k}_j^i$, player i best-responds to his belief over j's* behavioral level, *k_j^i. Here k_j^i is less than $\hat{k}_j^i$, because player i thinks that j best responds to his belief that i's bound is at $k_i^{ij} = 1$. Hence, i thinks that j's best response is $k_j^i = 2$, and i in turn best responds with $k_i = 3$.*

In Figure 3.b, $c_i^{ij} \in C^+(c_j^i)$, and therefore i believes that j views him as more sophisticated. Player i therefore expects j to play at his maximum bound, $k_j^i = \hat{k}_j^i = 3$. The best response is thus to play according to $k_i = 4$.

This example illustrates that when i believes that j is less sophisticated than he is himself, then i's choice depends on his second-order beliefs: i's level of play is lower when he believes that

[11] Note that setting $c_i^{ij} = c_i$, we obtain the case of eq. (5), where $k_j^i = \hat{k}_j^i$. This is so because, in that case, $\mathcal{K}(c_i^{ij}, v_i) = \hat{k}_i$, hence eq. (6) delivers $\hat{k}_i^{ij} = \hat{k}_j^i - 1$. By definition of $\hat{k}_j^i$, $\hat{k}_j^i - 1 < \hat{k}_i - 1$, hence eq. (7) implies $k_j^i = \hat{k}_i^{ij} + 1 = \hat{k}_j^i$. In fact, as the next example shows, the result from the previous subsection that $k_j^i = \hat{k}_j^i$ if $c_j^i \in C^-(c_i)$, requires only that $c_i^{ij} \in C^+(c_j^i)$; it is not necessary that $c_i^{ij} = c_i$.

j is underestimating i's sophistication. Extending this example would also show the following result. If i plays against a more sophisticated opponent and believes that the opponent believes that i is even more sophisticated, then i still plays according to his own bound, as he would if he thought that j had 'correct' beliefs about him. The higher order beliefs effects are thus *one-sided*, in the sense that they have behavioral implications only when referred to less sophisticated opponents.

Proposition 3 (Higher Order Effects) *Let $t_i = (c_i, c_j^i, c_i^{ij})$ be a* second-order type*, and let $v_i = v_j$.*

1. *Suppose $c_j^i \in C^-(c_i)$. For any $c_i^{ij} \in C^+(c_j^i)$, $k_i = \hat{k}_j^i + 1$. For $c_i^{ij} \in C^-(c_j^i)$, k_i (weakly) decreases as c_i^{ij} becomes less sophisticated.*
2. *Suppose $c_j^i \in C^+(c_i)$. For any $c_i^{ij} \in C^+(c_i)$ (rather than $c_i^{ij} \in C^+(c_j^i)$), $k_i = \hat{k}_i$. For $c_i^{ij} \in C^-(c_i)$, k_i (weakly) decreases as c_i^{ij} becomes less sophisticated.*

Testable Predictions for the 11-20 Game. The following proposition derives the predictions that will be tested in the experiment of Section 3. We emphasize that this proposition follows from Conditions 1 and 2 only, which (as discussed in Section 2.2.2) entail minimal restrictions on the shape of the functional forms. This proposition therefore allows us to test the 'detail free' implications of the cost-benefit approach. The proposition also shows that, even in its 'detail free' specification, our model delivers a rich set of testable predictions and provides a clear framework for the experimental design.

Proposition 4 *Consider the 11-20 game, with the extra reward for choosing the number exactly one below the opponent's parameterized by x. Under Conditions 1 and 2, for any i whose hierarchies of beliefs are described by second-order types, the following holds:*

1. ***Changing Incentives:*** *For any c_i, c_j^i, c_i^{ij}, the number chosen by player i is (weakly) decreasing in x.*
2. ***Changing Beliefs:*** *For any x and for any c_i and c_i^{ij}, the number chosen by player i (weakly) decreases as c_j^i becomes more sophisticated. Moreover, if $c_i^{ij} = c_i$, then i's cognitive bound is binding if he regards his opponent as more sophisticated (that is, $\hat{k}_i = k_i$ if $c_j^i \in C^+(c_i)$), not necessarily otherwise.*
3. ***Higher Order Beliefs Matter, but their Effects are One-Sided:*** *For any x and for any c_i and c_j^i, the number chosen by player i (weakly) decreases as c_i^{ij} becomes more sophisticated. Moreover:*
   - (a) *If i regards the opponent as less sophisticated ($c_j^i \in C^-(c_i)$), for any x, the number chosen by player i is constant in c_i^{ij}, as long as $c_i^{ij} \in C^+(c_j^i)$. For $c_i^{ij} \in C^-(c_j^i)$, the number decreases as c_i^{ij} gets more sophisticated.*

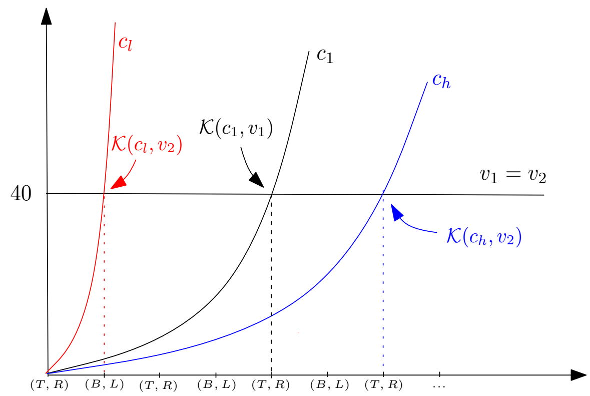


Figure 4: CTS with non-degenerate beliefs.

(b) If i regards the opponent as more sophisticated ($c_j^i \in C^+(c_i)$), for any x, the number chosen by player i instead is constant in c_i^{ij} as long as $c_i^{ij} \in C^+(c_i)$ (rather than $c_i^{ij} \in C^+(c_j^i)$). For $c_i^{ij} \in C^-(c_i)$, the number increases as c_i^{ij} gets less sophisticated.

2.3.2 General Model: Illustration and Discussion

The previous subsection has focused on cognitive type spaces with degenerate beliefs β_{t_i} for all types and players, allowing a type's first and second order beliefs to be written as cost functions c_j^i and c_i^{ij}, respectively. With non-degenerate beliefs, such cost functions must be replaced by distributions over cost functions. Nonetheless, Definition 5 implies that the recursion that characterizes players' play is essentially the same, as the following example illustrates:

Example 4 *Consider the coordination game provided at the start of this section (game B), with $x = 40$. Suppose, for simplicity, that the value of reasoning is constant and equal to 40 for both players, let the anchor be $a^0 = (T, R)$, the cognitive type space be as in Figure 4, with $T_1 = \{c_1\}$, $T_2 = \{c_l, c_h\}$, and let q denote the probability of $c_l \in \Delta(T_2)$.*

In words, the logic of the model is the following: The least sophisticated type, c_l, plays according to his cognitive bound, $\mathcal{K}(c_l, v_2) = 1$, which corresponds to action L. The only type of player 1 has depth $\mathcal{K}(c_1, v_1) = 4$. This corresponds to action $a_1^4 = T$ in his path of reasoning, but it need not be the action that he plays: if he places probability $q > 1/2$ of facing a c_l player 2, then he best responds to c_l's action, and hence plays as a level-2, with associated action B. Otherwise, c_1 would play according to his own bound, as a level-4. The more sophisticated c_h understands the behavior of the other types, and best responds to c_1's behavior. If $q > 1/2$, c_h anticipates that c_1 would play as a level-2, and behave as a level-3; otherwise, c_h anticipates that c_1 plays as a level-4, and behave as a level-5.

Applying the recursion (3) in Definition 5 results in the argument above. To see this, note that the least sophisticated type c_l reaches his cognitive bound at 1, $\alpha_2^k(c_l) = a_2^1 = L$ for all $k \geq 1$, hence $\hat{a}_2(c_l) = a_2^1 = L$. For the other types, c_1 and c_h, the recursion yields, respectively, $\alpha_1^1(c_1) = a_1^1 = B$ and $\alpha_2^1(c_h) = a_2^1 = L$ for the first step, and $\alpha_1^2(c_1) = BR(q\alpha_2^1(c_l) + (1-$

$q)\alpha_2^1(c_h)) = T$ and $\alpha_2^2(c_h) = BR(\alpha_1^1(c_1)) = R$ for the second. Notice that, for the second step, the action of player 2 is not the same for type c_l and c_h. Hence, for subsequent steps, the action of player 1 - and therefore, the action of player 2 with type c_h- depend on q. Specifically, for c_1, the third step yields $\alpha_1^3(c_1) = BR(q \cdot L + (1-q)\alpha_2^2(c_h))$, which is equal to $a_1^3 = T$ if $q > 1/2$, and B otherwise. In the first case, the iteration stays constant for type c_1. Otherwise, it follows the path of reasoning. As for the most sophisticated type c_h, each step corresponds to the best response to c_1, to whom he attaches probability one.

As this example illustrates, the recursion in Definition 5 coincides with the path of reasoning as long as $k < \min_{j\in N, t_j \in T_j} \mathcal{K}(c_{t_j}, v_j)$. That is, the path of play follows the path of reasoning as long as no player has reached his cognitive bound. For types that have the lowest depth of reasoning, given their incentives, the path of play becomes constant for iterations above their cognitive bound. Therefore, they play according to their own bound: $\hat{a}_i(t_i) = a_i^{\hat{k}_{t_i}}$ when $\hat{k}_{t_i} = \min_{j\in N, t_j \in T_j} \mathcal{K}(c_{t_j}, v_j)$. Recursively, the path of play also departs from the path of reasoning for types that place sufficiently high probability on types whose own path of play differs from their path of reasoning. This could be either because they have reached their bound, or (recursively) because they believe the opponent has, and so forth.

In addition to assuming degenerate beliefs, the simplified model of Section 2.3.1 makes the important restriction that first and second order beliefs suffice to pin down the entire hierarchy of beliefs. In general, belief hierarchies can be more complicated, and lead to more complex patterns of behavior. However, the logic of the recursion captures the idea that even though a cognitive type space entails fully formed hierarchies of beliefs over cost functions, higher order beliefs effects are bounded by players' depth of reasoning, which is consistent with the 'one-sidedness' results of Propositions 3 and 4. Therefore, given the recursion in Definition 5, cognitive type spaces are an effective device for modeling the limited effects of higher order beliefs in the presence of bounded depth of reasoning.

2.4 Related Models

Within the literature on level-k reasoning, the closest model to ours is Strzalecki's (2010), which also separates depth of reasoning from beliefs and behavior. In his model, the depth of reasoning is given, and each type's beliefs are concentrated on types with lower depth of reasoning. Our model endogenizes the depth of reasoning and allows players to believe that their opponent is more sophisticated. We note that this notion is natural in our setting, thereby resolving a well-known conceptual difficulty of the level-k approach. Strzalecki's model can be nested in ours, letting cost functions be zero and then infinite at some fixed k, and letting beliefs be concentrated on less sophisticated types. In that case, the recursion in Definition 5 has the same behavioral implications as Strzalecki's equilibrium concept. If such beliefs are further assumed to be equal to the correct distribution of types with lower depth of reasoning, then the CH model of Camerer et al. (2004) obtains as a further special case. The separation of beliefs from behavior thus enables us to accommodate both the CH model and the models in which level-k types best respond to level-$(k-1)$ (e.g., Nagel (1995), Costa-Gomes and Crawford (2006), Crawford and Iriberri (2007), and others). This separation is particularly important in the context of initial response games, because the assumption that agents have correct beliefs about the distribution of depths of reasoning is not justified on the grounds of introspection alone. The model nonetheless allows for that possibility, and in particular we impose it as an identification restriction for the calibration exercise in Section 5.

The general notion that players follow a cost-benefit analysis is present in the language of Camerer, Ho and Chong (2004), but not in their model itself, as players' cognitive types remain exogenous. A recent paper by Choi (2012) extends Camerer, Ho and Chong's (2004) model by letting cognitive types result from an optimal choice. This optimization serves to provide identification restrictions to estimate the distribution of types across different environments, and is motivated by an evolutionary argument. In contrast, our goal is to provide an explicit model of reasoning, in which players think about both the game and about the reasoning process of their opponents. The objectives and modeling choices are therefore distant.

Gabaix (2012) also proposes a framework in which the accuracy of players' beliefs about the opponents' behavior is determined by a cost-benefit analysis. At a conceptual level, the main goal of Gabaix (2012) is to provide an equilibrium concept that allows players to have both incorrect beliefs as well as to respond non-optimally. Our model focuses on agents' cognitive limitations in reasoning about the opponents, and ignores the important but orthogonal issue of limitations in computing best responses. The 'noisy introspection' model of Goeree and Holt (2004) extends the level-k approach introducing non-optimal responses in a non-equilibrium model. Introducing noisy responses in our model of endogenous depth of reasoning is an interesting direction for future research.

From a broader perspective, our approach can be cast within the research agenda on rational inattention, which also endogenizes individuals' limited understanding of the environment through a cost-benefit approach.[12] This literature has thus far focused on non-strategic problems. Strategic settings raise specific complications, particularly due to the interaction between individuals' understanding, their beliefs and their higher order beliefs. From a conceptual viewpoint, the literature on unawareness is also related (for a thorough survey, see Schipper (2014)). While models of unawareness in strategic settings have a different focus, our framework can be viewed as endogenizing the awareness of the opponents' best responses. Exploring the formal connections would be another interesting avenue for future research.

3 Experimental Design

The experiment tests the key implications of the detail-free model, on the manner in which behavior is affected by the incentives to reason, the beliefs about the opponents and the higher order beliefs. The experimental design matches closely the theoretical setting, in which each

[12]The classical reference for rational inattention is Sims (2003), which spurred a large literature. A recent paper closely related to our work is Caplin and Dean (2013).

change occurs in isolation. In particular, testing the effects of changing the incentives to reason requires that they vary but that the costs of reasoning and belief hierarchies remain fixed. Similarly, we test the effects of changing only beliefs over the opponent, and the effects of changing only beliefs over the opponent's beliefs. We discuss below how the experiment achieves these objectives. The precise mapping from the theoretical predictions, stated in Proposition 4, and the treatments of the experiment, presented in Section 3.1 and 3.2, are summarized in Section 3.3.

Throughout the treatments, the baseline game remains the modified 11-20 game discussed in Section 2:

> The subjects are matched in pairs. Each subject enters an (integer) number between 11 and 20, and always receives that amount in tokens. If he chooses *exactly one less* than his opponent, then he receives an extra 20 tokens. If they both choose the same number, then they both receive an extra 10 tokens.

This game is a variation of Arad and Rubinstein's (2012) '11-20' game, the distinction being that the original version does not include the extra reward in case of a tie. As in the original 11-20 game, the best response to 20 (or to the uniform distribution) is 19, the best response to 19 is 18, and so forth. But with the extra reward in case of tie, the best response to 11 is 11, and not 20, as is the case in the original 11-20 game. Thus, our modification breaks the cycle in the chain of best responses. As discussed in Section 3.4, this game presents several advantages for our purposes.

The subjects of the experiment were 120 undergraduate students from different departments at the Universitat Pompeu Fabra (UPF), in Barcelona. Each subject played twice every treatment described in Sections 3.1 and 3.2, and summarized in Table 1. They also played a subset of the additional treatments described in Appendix E. These treatments are all based on the modified 11-20 game. We provide the exact sequences of treatments used in Appendix A.2.

Each subject was anonymously paired with a new opponent after every iteration of the game. To focus on initial responses and to avoid learning from taking place, the subjects did not observe their payoffs after their play. They only observed their earnings at the end of the session. Each token was worth five euro cents. Moreover, subjects were paid randomly, and therefore did not have any mechanism for hedging against risk by changing their actions.[13] Specifically, they were paid once for each set of six iterations, and the randomization occurred inside this set (once at random for the first six iterations, once for the next six iterations, and so forth). As an additional control for order effects, the order of treatments was randomized. Furthermore, since subjects played the same treatments twice during a session, we can compare play for each treatment through equality of distribution tests. Lastly, subjects received no

[13] These methods are standard in the literature that focuses on 'initial responses', where the classical equilibrium approach is hard to justify. See, for instance, Stahl and Wilson (1994, 1995), Costa-Gomes, Crawford and Broseta (2001) and Costa-Gomes and Crawford (2006). For an experimental study of equilibrium in a related game, see Capra, Goeree, Gomez and Holt (1999).

Treatment	Own label	Opponent's label	Own payoffs	Opponent's payoffs	Replacement of opponent's opponent
Homogeneous [A]	I (II)	I (II)	Low	Low	No
Heterogeneous [B]	I (II)	II (I)	Low	Low	No
Replacement [C]	I (II)	II (I)	Low	Low	Yes
Homogeneous-high [A+]	I (II)	I (II)	High	High	No
Heterogeneous-high [B+]	I (II)	II (I)	High	High	No
Replacement-high [C+]	I (II)	II (I)	High	High	Yes

Table 1: Treatment summary: Label I refers to 'math and sciences' or to 'high' subjects, and label II refers to 'humanities' or to 'low' subjects. There are 120 subjects for each treatment (60 subjects for each classification).

information concerning their opponents' results. This serves to avoid that subjects focus on goals other than monetary incentives, such as defeating the opponent or winning for its own sake. The instructions of the experiment were given in Spanish; the English translation and the details on the pool of subjects, the earnings and the logistics of the experiment are in Appendix A.

3.1 Changing beliefs about the opponents

We consider two different classifications of subjects, an *exogenous classification* and an *endogenous classification*, each with 3 sessions of 20 subjects. In the exogenous classification, subjects are distinguished by their degree of study. Specifically, in each session of the experiment, 10 students are drawn from the field of humanities (humanities, human resources, and translation), and 10 from math and sciences (math, computer science, electrical engineering, biology and economics). The subjects are aware of their own classification when beginning the experiment, and are labeled as 'humanities' or 'math and sciences'. In the endogenous classification, there is no restriction on the pool of subjects. Moreover, the subjects are not informed about the field of study of the other players. Before playing the game, however, they are required to take a test of our design. Based on their performance on this test, each student is either labeled as 'high' or 'low', and is shown his own label before playing the game. We defer the description of this test to Section 3.4.1 (see also Appendix B).

These classifications allow us to change subjects' beliefs about their opponents. In each treatment, the subjects are given information concerning their opponents. They play the baseline game against someone from their own label (homogeneous treatment [A]) and against someone from the other label (heterogeneous treatment [B]). For instance, for the exogenous classification, a student from math and sciences (resp., humanities), is told in homogeneous treatment [A] that his opponent is a student from math and sciences (humanities) as well. In heterogeneous treatment [B], he is told that the opponent is a student from humanities (math and sciences). Identical instructions are used for the endogenous classification, but with 'high' and 'low' instead of 'math and sciences' and 'humanities', respectively.

Suppose, for the sake of illustration, that there are two cognitive categories of subjects, consisting of those with higher game theoretical sophistication and those with lower game theoretical sophistication. For our purposes, it suffices that players *believe* that there is a meaningful difference between these two categories, and that the labels we use in the two classifications are informative about their opponents' category (see Section 3.4). We hypothesize that subjects in the exogenous classification associate the 'maths and sciences' label and 'humanities' label with higher and lower game theoretical sophistication, respectively, and that subjects in the endogenous classification associate the 'high' and 'low' labels with higher and lower game theoretical sophistication, respectively.[14] Then, when playing homogeneous treatment [A] compared to heterogeneous treatment [B], subjects change from believing they are playing against one category of player to another.

Treatments [A] and [B] are designed to test whether the behavior of the subjects varies with the sophistication of the opponent. From these treatments, however, it is difficult to establish whether the subjects are aware that their opponents' beliefs may affect their choices. The next treatment is designed to test whether the subjects believe that (or are aware of) the behavior of their opponents also changes when they face opponents of different levels of sophistication. To test for these higher order beliefs effects, we administer replacement treatment [C]. A 'math and sciences' subject, for instance, is given the following instructions: "[...] two students from humanities play against each other. You play against the number that one of them has picked." The reasons for using this exact wording are discussed in Section 3.4.2.

3.2 Changing incentives

We next consider a second dimension that would entail a change in players' chosen actions, according to our framework. In particular, we aim to test the central premise of our theoretical model, that players may perform more rounds of introspection if they are given more incentives to do so. To do this, we change the baseline game in the following way: rather than winning an extra 20 tokens for choosing the number precisely one below the opponent, subjects win an extra *80* tokens. The rest of the game payoffs remain the same. It is immediate that this change does not affect the path of reasoning, irrespective of whether the level-0 is specified as 20 or as the uniform distribution. It only increases the rewards for players who stop at the 'correct' round of reasoning. In the context of our theoretical model, this game is in the same cognitive equivalence class as the baseline game, and so the costs of reasoning are identical.

We consider three treatments for this 'high payoff game': homogeneous treatment [A+], heterogeneous treatment [B+], and replacement treatment [C+]. These treatments are equivalent to treatments [A], [B] and [C], respectively, but with the increased reward for undercutting of 80 tokens. We then compare an agent's play under lower payoffs to his play under higher payoffs, by comparing [A] to [A+], [B] to [B+] and [C] to [C+]. We also compare treatments [A+], [B+] and [C+] in an analogous way to the comparison between treatments [A], [B] and [C].

[14]The only assumption that is needed is that the two labels are perceived as different in terms of strategic sophistication. Since, however, the data reveal that label I subjects ('math and sciences' and 'high') are regarded as more sophisticated, we maintain this assumption in this discussion.

This concludes our discussion of the main treatments. The next Section explains how these treatments relate to the theoretical model, and presents the theoretical predictions for the experiment. These predictions are summarized in Table 2.

3.3 Theoretical Predictions for the Experiment

We describe next how we apply our theory to the treatments of the experiment. Recall that we use the terminology 'label I' (resp., 'label II') to refer to the 'high score' ('low score') subjects in the endogenous classification or to the 'math and sciences' ('humanities') subjects for the exogenous. Accordingly, we introduce notation $l_i = \{I, II\}$ to refer to individual i's label. For simplicity, we only consider the second-order types discussed in Section 2.3.1.

First, we assume that an individual's cost of reasoning, c_i, remains constant throughout all treatments. This is consistent with the cognitive equivalence of the games used in the low and high payoff treatments. We also assume that i's first-order beliefs, c_j^i, only depend on the label of the opponent, and that his second order beliefs c_i^{ij} only depend on the label of the opponent's opponent (which is i's own label, except in the 'replacement' treatments, $[C]$ and $[C+]$). This implies that an individual is identified by his label l_i, his cost c_i and first and second order beliefs in treatment $[X]$, denoted by beliefs $c_j^{i,[X]}$ and $c_i^{ij,[X]}$ (for $X = A, B, C$), which satisfy the following:

E.1: For all i: $c_j^{i,[B]} = c_j^{i,[C]}$, $c_i^{ij,[A]} = c_i^{ij,[B]}$ and for all $X = A, B, C$, $c_j^{i,[X]} = c_j^{i,[X+]}$and $c_i^{ij,[X]} = c_i^{ij,[X+]}$.

We also assume that individuals commonly believe that label I players are more sophisticated than label II. Formally:

E.2: For label I individuals: if $l_i = I$, $c_j^{i,[B]} \in C^-(c_j^{i,[A]})$, $c_i^{ij,[C]} \in C^-(c_i^{ij,[B]})$; For label II individuals: if $l_i = II$, $c_j^{i,[A]} \in C^-(c_j^{i,[B]})$, $c_i^{ij,[B]} \in C^-(c_i^{ij,[C]})$.

Finally, we assume that label II individuals always regard label I's as more sophisticated then they are:

E.3: $c_j^{i,[B]} \in C^+(c_i)$ whenever $l_i = II$.

Under E.1 and E.2, for any $X = A, B, C$, the only change between treatment $[X]$ and $[X+]$ is in the payoffs x. These comparisons therefore allow us to test the implications of part 1 of Proposition 4. Treatments $[A]$ and $[B]$ (or $[A+]$ and $[B+]$) instead only differ in i's first order beliefs, they thus serve to test part 2 of Proposition 4. Finally, treatments $[B]$ and $[C]$ (or $[B+]$ and $[C+]$) only differ in i's second order beliefs. Their comparison therefore addresses the third part of Proposition 4.

More specifically, let F_X^l denote the cumulative distribution of actions $a \in \{11, ..., 20\}$ in treatment X for label $l \in \{I, II\}$, and denote by $\succsim$ the first order stochastic dominance relation.[15] Proposition 4 immediately implies the following results, summarized in Table 2:

[15] Given two cumulative distributions $F(x)$ and $G(x)$, we say that F (weakly) first order stochastically dominates G, written $F \succsim G$, if $F(x) \leq G(x)$ for every x.

Labels	Changing beliefs (low payoffs)	Changing beliefs (high payoffs)	Changing payoffs
$l_i = I$	$F_C \succsim F_B \succsim F_A$	$F_{C+} \succsim F_{B+} \succsim F_{A+}$	$F_X \succsim F_{X+}$ for $X = A, B, C$
$l_i = II$	$F_A \succsim F_B$; $F_B \approx F_C$	$F_{A+} \succsim F_{B+}$; $F_{B+} \approx F_{C+}$	$F_X \succsim F_{X+}$ for $X = A, B, C$

Table 2: Summary of the theoretical predictions of the relations between the distribution of actions in different treatments.

Proposition 5 *For any distribution over individuals that satisfy the restrictions in E.1-3, under the maintained assumptions of the detail-free model (Section 2), the following holds: (i) For any $X = A, B, C$ and $l = I, II$, $F^l_X \succsim F^l_{X+}$; (ii) $F^I_C \succsim F^I_B \succsim F^I_A$; (iii) $F^{II}_A \succsim F^{II}_B \approx F^{II}_C$.*

3.4 Experimental Design: Discussion

3.4.1 Designing the Group Classification: Demarcation and Focality.

In order to vary subjects' beliefs about the opponents, we divide the pool of subjects into two labeled groups. We then change subjects' beliefs about the opponents by changing the opponent's group in the different treatments. To effectively implement the theoretical proposition that we test, and in particular conditions E.2-3 in Section 3.3, these labels must satisfy two properties. The first is *demarcation*: the groups in the classifications must be perceived as being sufficiently distinct from each other in terms of their strategic sophistication. The second is *focality*: since subjects' behavior depends not only on their beliefs but also on their beliefs about their opponents' beliefs, it is important that the two groups share sufficient agreement about the way they differ. That is, they should 'commonly agree' over their relative sophistication. The two classifications we consider have been chosen to guarantee that these properties hold.

The Exogenous Classification. The exogenous classification exploits the intuitive, albeit vague, view that 'math and sciences' students are regarded as more accustomed to numerical reasoning than 'humanities' students. Furthermore, the specific degrees of study used to populate the 'math and sciences' group are commonly viewed as being the most selective degrees at UPF, and require the highest entry marks.[16] We would therefore expect the subjects to believe the 'math and sciences' group to be comparatively more sophisticated in game theoretical reasoning than the 'humanities' group. However, the subjects are not primed into shaping specific beliefs about either particular group.

The Endogenous Classification and the Test. Since the exogenous classification is based on labels that are salient for the subjects in the pool, it can be expected to ensure both

[16] These views emerged from informal conversations with students. They are confirmed by the admission scores, used to select the students admitted in the various fields. These scores can be found at: http://www.elpais.com/especial/universidades/titulaciones/universidad/universidad-pompeu-fabra/45/nota-corte/ .

demarcation and focality. But it does not allow us to fully control the agents' beliefs about these labels. For this reason, we also introduce the endogenous classification, where students are classified solely based on their performance in a test of our design. The goal of the test is thus twofold. It sorts subjects into two groups, and, by labeling the scores obtained by subjects as 'high' or 'low', the test itself forms the agents' beliefs over the content of these labels.

The main objective of the test is to convince subjects that the result of the test is informative about their opponents' game theoretical sophistication. To do so, we ensure that our test questions appear difficult to solve, and that subjects would be likely to infer that an individual of higher sophistication would respond better to the questions.[17]

The cognitive test takes roughly thirty minutes to complete, and consists of three questions. These are all single-agent questions, in that they are not pitted against each other. Rather, subjects are asked to provide the correct answer. Scores are assigned to the answers to each of these tasks, and a formula then takes a weighted average. Those who have a score above the median are labeled 'high', and the others are labeled 'low'. Subjects do not see their numerical grade, but they are told whether they are labeled 'high' or 'low'. (Details of the test are contained in Appendix B).

The three questions are as follows. In the first question, subjects have nine attempts to play a variation of the board game *mastermind*, in which the aim is to deduce a hidden pattern through a sequence of guesses. This game requires skill at logical inference to complete successfully. In the second question, they are given a typical *centipede game* of seven rounds. In the third game, the agents are given a lesser known *pirates game.* As with the centipede game, the pirates game can be solved by backward induction. The solution is more difficult to attain, however, in that it involves additional computation in determining how players' strategy profiles map into outcomes. For the latter two games, subjects are not asked how they would play; rather, they are asked how "infinitely sophisticated and rational agents, who each want to get as much money as possible" would play.

These three questions are challenging to most, and performing well is arguably informative of a subject's sophistication. More importantly, for our purposes, it seems fair to assume that the subjects believe the score to be a strong indicator of game theoretical sophistication. Arguably, this would be the case for subjects who recognize that the 11-20 game has a recursive pattern reminiscent of the structure of the problems in the test, which appears plausible for players who would perform at least one round of introspection. Furthermore, while the last two games are close enough to the 11-20 game to require a similar kind of analysis, they are sufficiently different that the feedback provided in the form of the 'high' or 'low' score should limit the learning about how to play the 11-20. It is precisely to avoid information leaking from the test to the proper experiment that we have not included the beauty contest, the traveler's dilemma or other closely related games.

[17] Another possible consequence of being assigned a 'high' or a 'low' label may impact an individual's self-esteem, which may in turn impact his subsequent performance in playing the modified 11-20 game. This concern, however, is tangential to the aim of our experiment. Our objective is not to identify the number of rounds of introspection a subject performs in a single game, but how his actions vary across treatments.

3.4.2 Testing for Effects of Higher Order Beliefs

The objective of treatment [C] is to test for the degree to which the subjects have a well-formed model of their opponents' reasoning, and whether it is consistent with our theoretical predictions. The precise wording of treatment [C] is designed to pin down the entire hierarchy of beliefs, as described in Section 2.3.1. For instance, the full description that a math and sciences student is given concerning his opponent in treatment [C] is: "[...] two students from humanities play against each other. You play against the number that one of them has picked." It is therefore clear that he is playing a humanities playing a humanities subject, who himself is playing a humanities subject, and so forth. Any ambiguity that allowed players to believe that one of them believes (at some high level) that one of them is a student from math and sciences could result in a sophisticated subject behaving as a less sophisticated one, invalidating the identification of the types.

3.4.3 Choice of the Baseline Game

As argued by Arad and Rubinstein (2012), the 11-20 game presents a number of advantages in the study of sequential reasoning, which are inherited by our modified version. We recall here the most relevant to our purposes. First, using sequential reasoning is natural, as there are no other obvious focal ways of approaching the game. The competing alternative of guessing the unique pure-strategy equilibrium seems far from self-evident, and would be difficult to see without going through at least a few steps of iterated reasoning. Secondly, the specification of the anchor is intuitively appealing and unambiguous, since choosing 20 is natural for an iterative reasoning process. Moreover, it is the unique best choice for a player who ignores all strategic considerations. Thirdly, there is robustness to the a^0, in that the choice of 19 would be the best response for a wide range of anchors, including the uniform distribution over the possible actions. Lastly, best-responding to any action is simple. Since we do not aim to capture cognitive limitations due to computational complexity, having a simple set of best responses is preferable in this case.

In addition to these points, our modification of the game leads to another useful feature. Since the best response to 11 is 11, and not 20, our modification breaks the cycle in the chain of best responses. It is crucial that these cycles be avoided here, since one of the main hypotheses that we aim to test is whether an increase in players' incentives would shift the distribution of play. This hypothesis could not be falsified in the presence of such cycles.[18]

While the modified 11-20 game is particularly suitable to our purposes, our model applies to a wide spectrum of games. We discuss here some related games to which our framework appears especially relevant.

Basu's (1994) traveler's dilemma is perhaps the closest to ours, and is one of the games discussed in Section 5. In this two-player game, each agent reports a number between 2 and

[18] Our theoretical predictions on the shift of the distribution do not depend on assumptions of degenerate beliefs; as discussed in Section 2, our model allows for non-degenerate beliefs. Noise in the path of reasoning, in the spirit of Goeree and Holt (2004), can be introduced as well.

100, and both players receive the lowest of the two numbers chosen. In addition, if the players report different numbers, then the one who reports a higher number pays a penalty of 2, and the one with the lowest receives a reward of 2. This game shares appealing features of the modified 11-20 game. The main difference is that it is sufficient in the traveler's dilemma to undercut the opponent to receive the additional reward, rather than to choose exactly the right action. The modified 11-20 game therefore leads to a more precise mapping between the agent's action and his beliefs. Moreover, if agents had social preferences such as altruism or fairness, it would not be an issue in the modified 11-20 game. This is because, independent of the agents' preferences over the final outcomes, the optimal choice would still require players to identify their opponents' action.

A central position in the literature on level-k is occupied by Nagel's (1995) *beauty contest* game, also known as the *p-guessing game.*[19] In this game, n players are asked to report a number between 0 and 100, and the player whose number is closest to a fraction $p \in (0,1)$ of the average report wins a monetary prize. A remarkable finding of the literature is that players' reports exhibit a regular pattern concentrated around specific numbers, and that the position of these spikes shifts down as p decreases. This evidence has been interpreted as suggesting that players approach games of this kind by thinking in steps, which has been a key motivation to the development of level-k theories. Importantly, the large strategy space of this game allows testing whether players reason according to some form of iterated reasoning.[20] Our goal, however, is different: here we maintain that players follow an iterated reasoning procedure, and we ask whether their depth of reasoning varies with their beliefs about the sophistication of the opponents, and with incentives to reason more deeply about the game. A number of features of the beauty contest makes it less suitable than the modified 11-20 game for this task. For instance, level-k reasoning is not necessarily the only 'focal' form of reasoning, and, for related reasons, the beauty contest does not present an obvious specification of the level-0 action.[21] This last point is particularly problematic because the level-1 action is highly sensitive to the specification of level-0 in this game. Furthermore, a natural setting for our experiment is one with few players, as it allows for greater ease in changing beliefs over opponents, and opponents' opponents. But with a small number of players, the best response function in the p-guessing game is difficult to compute. A player must take into account the impact of his own number on the average, which is not as immediate a calculation as computing the best response in our game (see Grosskopf and Nagel 2008). Notwithstanding these factors, we expect the comparative statics predictions of our theoretical model to hold for the beauty contest as well.

[19] Variations of this game have been studied, among others, by Ho, Camerer and Weigelt (1998) and Bosch-Domènech, Garcia-Montalvo, Nagel and Satorra (2002).

[20] For studies that focus more directly on the cognitive process itself, see Agranov, Caplin and Tergiman (2012), and the recent works by Bhatt and Camerer (2005), Coricelli and Nagel (2009), and Bhatt, Lohrenz, Camerer and Montague (2010), which use fMRI methods and find further support for level-k models. See also Georganas, Healy and Weber (2012), Gill and Prowse (2013) and Fehr and Huck (2013) for analyses of cognitive ability in strategic settings.

[21] For a thorough description of these different thought processes, see Bosch-Domènech, García-Montalvo, Nagel and Satorra (2002).

In another important game, introduced by Costa-Gomes and Crawford (2006, CGC), the players' objective is to guess a target that depends on their opponent's guess multiplied by a constant. Because of the large strategy space, this game inherits the appealing properties of the beauty contest. Like our game, however, it is a two-player game with simple best response functions. It further allows a separation between different types of reasoning processes, such as level-k reasoning and iterated dominance, which is not allowed by the standard beauty contest. More importantly, subjects in CGC play a sequence of games in which both the strategy space and the target are varied. Such sequences of responses yield strategic 'fingerprints' that allow CGC to identify individuals' types of reasoning. As discussed earlier, our objective is distinct, as we do not aim to separate level-k from other forms of reasoning.

4 Experimental Results

We present, for brevity, only the experimental results for the grouped exogenous and endogenous classifications.[22] We pool the label I subjects ('math and sciences' for exogenous treatments and 'high' for endogenous treatments), and we pool the label II subjects ('humanities' for exogenous treatments and 'low' for endogenous treatments). Recall that we do not take these labels to indicate actual game theoretical sophistication, but of *perceived* sophistication by the subjects. Moreover, we present the results by pooling together the treatments when they are repeated. For these repetitions, our pooling is justified by tests for equality of distribution. We analyze first the results when subjects' payoffs are changed, followed by the results when their beliefs over opponents are varied. Overall, the results appear to confirm the theory. We consider Wilcoxon-signed-rank tests, and to further test our first order stochastic dominance relations, we use the method introduced by Davidson and Duclos (2000) (henceforth, DD). We also conduct regressions, and find that most of the relevant coefficients are significant with the sign expected. This lends support to our model, and to the general claim that beliefs and incentives affect the individuals' depth of reasoning and their behavior.

In the (random-effects) ordinary least squares estimations (OLS) that follow, we regress, for each label, the outcome on a dummy for the treatments, and another for the classification (endogenous or exogenous). The latter is never significant. To control for 'feedback-free' learning, we exploit two factors. First, we use randomization of treatments, particularly within [A], [B] and [C], and within [A+], [B+] and [C+]. Second, we exploit the repetition of treatments to do equality of distribution tests for the same treatment. For instance, each agent plays treatment [A] twice with other treatments in between, and we test for equality of distribution between the first and the second time treatment [A] occurs. The results of these tests are highly suggestive that the distributions are equal. As an added robustness check in comparing low to high payoffs in the OLS estimations, we also control for the round at which they are played, and find that it is never significant.

[22] The figures for the separate classifications are consistent with the results for the grouped classifications. They are available upon request, as are other supplementary materials.

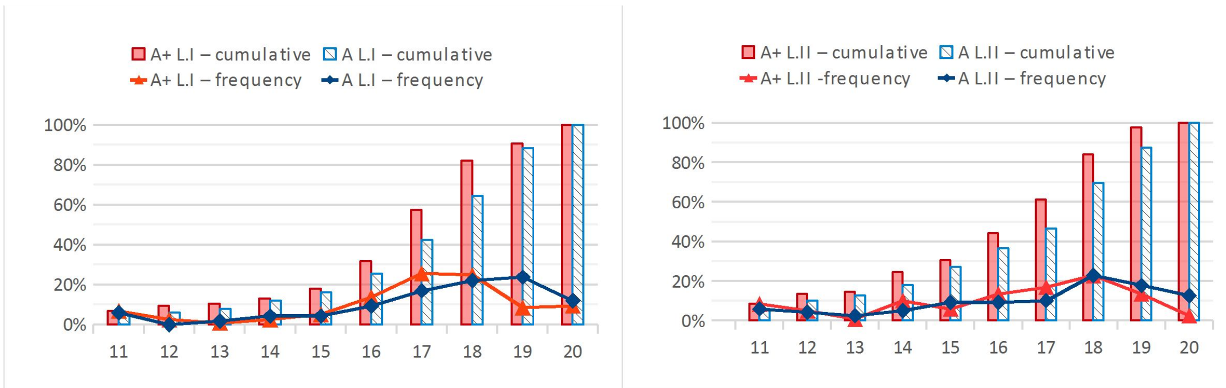
All regressions and statistical tests are in Appendix C. The OLS regressions are in Table 5 of Appendix C and the Wilcoxon-signed-rank tests for changes in payoffs and beliefs over opponents are in Table 6 and Table 7, respectively. The DD tests are in Tables 8 and 9.

4.1 Changing incentives

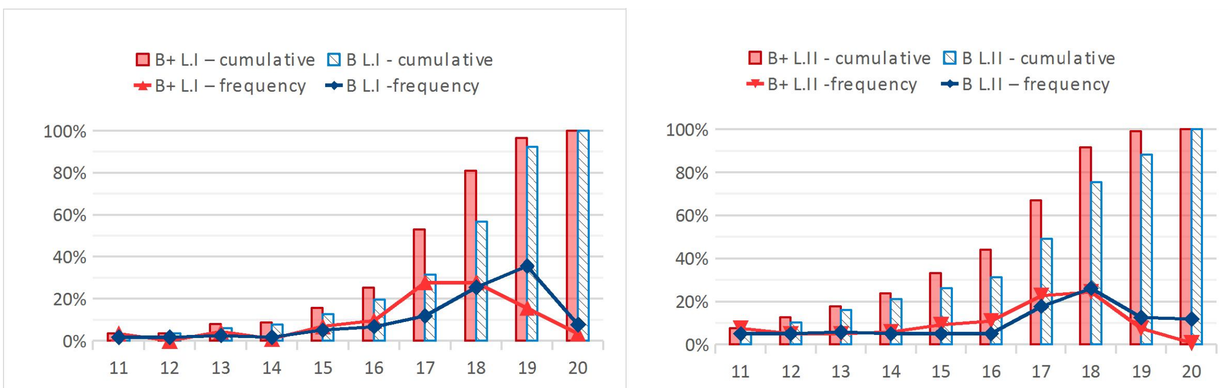
As the value of introspection increases for players and their opponents, the model predicts that they would choose actions associated with higher k's. Specifically, comparing treatments across different marginal values of payoffs, $F_A \succeq F_{A+}$, $F_B \succeq F_{B+}$ and $F_C \succeq F_{C+}$. These implications hold for both label I and label II subjects. Beginning with label I, it is clear from Figure 5.a (left) that the empirical distribution [A] stochastically dominates [A+] everywhere. Furthermore, distribution [B] stochastically dominates [B+] everywhere, and [C] clearly stochastically dominates [C+] everywhere (Figures 5.b and 5.c). Using a DD test for each comparison, we find that these results are indeed consistent with first order stochastic dominance, as shown by all the signs of the statistics. These results are therefore consistent with our theoretical predictions. Conducting an OLS regression, we find that the coefficients are highly significant ($< 1\%$) for distributions [A] compared to [A+], [B] to [B+] and [C] to [C+], and of the correct sign. The Wilcoxon-signed-rank is highly significant ($< 1\%$), for all of these comparisons of distribution as well.

These findings are consistent with the model, and with the view that agents perform more rounds of reasoning if the incentives are increased. These results also indicate that changing from an extra 20 tokens to an extra 80 tokens determines a large enough shift in the value function that it leads agents to increase their level of reasoning. The graphs in Figure 5 are suggestive of the idea that there is a shift in the distribution. We note as well that level-1 and level-2 play is modal for treatments [A], [B] and [C], while level-2 and level-3 play is modal for [A+], [B+] and [C+]. The means of the distributions change from 17.3, 17.7 and 18.0 for treatments [A], [B] and [C], respectively, to $16.8, 17.1$ and 16.9 for treatments [A+], [B+] and [C+], respectively.

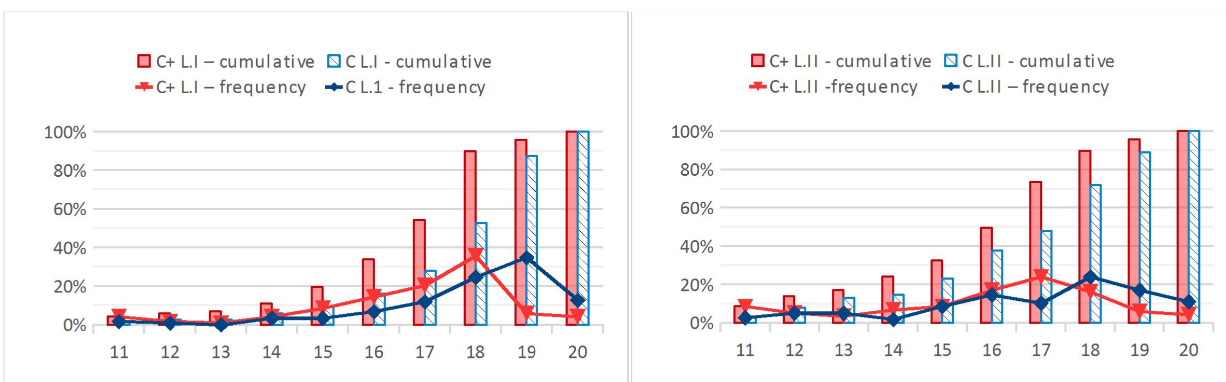
For label II, the stochastic dominance relationships hold everywhere for all three comparisons, [A] to [A+], [B] to [B+] and [C] to [C+], as shown in Figure 5 (right). All the signs of the statistics of the DD tests also confirm that these results are consistent with stochastic dominance, and the coefficients from the OLS regression are significant ($< 5\%$) for distributions [A] compared to [A+] and highly significant ($< 1\%$) for distributions [B] compared to [B+] and [C] to [C+]. These coefficients are of the correct sign. The Wilcoxon-signed-rank is highly significant ($< 1\%$) for all of these comparisons. As with label I, the figures for label II are suggestive of a shift in the distribution. Here as well level-1 and level-2 play is modal for treatments [A] and [C] (level-2 and level-3 play is modal for [B]), while level-2 and level-3 play is modal for [A+],[B+] and [C+]. The means of the distributions change from 16.9, 16.8 and 16.9 for treatments [A], [B] and [C], respectively, to $15.6, 15.6$ and 15.5 for treatments [A+], [B+] and [C+], respectively.



(a) : A and A+



(b) : B and B+



(c) : C and C+

Figure 5: Changing Payoffs, label I (left) and label II (right)

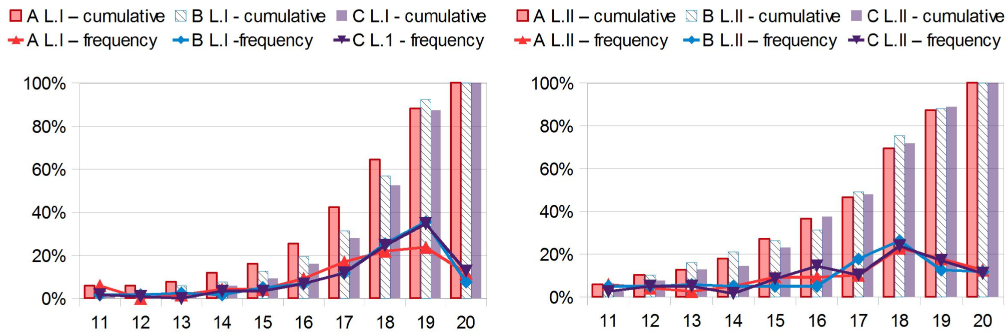


Figure 6: Treatments [A], [B] and [C] for label I (left) and II (right).

4.2 Changing beliefs about the opponents

Consider the comparison between homogeneous treatment [A], heterogeneous treatment [B] and replacement treatment [C]. According to the theoretical model, $F_C \succsim F_B \succsim F_A$ for label I players. This result seems consistent with the data displayed in Figure 6. Distribution [C] clearly stochastically dominates [B] everywhere, and [B] stochastically dominates [A] nearly everywhere.[23] We also note that [C] clearly stochastically dominates [A] everywhere.

Note that the theoretical model does not predict *strict* stochastic dominance relations $F_C \succ F_B$ or $F_B \succ F_A$. That is, it allows for individuals' beliefs over their relative sophistication to be such that they would play the same against lower or higher levels of sophistication. The distinct pattern that emerges from Figure 6 indicates that label I individuals view the cost function associated with label II as sufficiently far from their own to induce a difference in their chosen action. We also note that this pattern could not be explained by existing models of level-k reasoning, as they do not endogenize that level of play may vary even if the payoffs of the game remain constant.

The OLS estimates comparing [A] to [B] are significant at $< 10\%$, and the estimates comparing [A] to [C] are highly significant ($< 1\%$). The estimates comparing [B] to [C], however, are not significant. Figure 6 reveals that distributions [B] and [C] remain very close to each other, and so the lack of significance is not surprising.

Turning next to label II players, the model predicts $F_A \succsim F_B \approx F_C$. Here, no clear difference emerges from Figure 6 between the three cumulative distributions. Conducting Wilcoxon-signed-rank equality of distribution tests confirms the visual intuition, and the OLS estimates are not significant for any of the comparisons of [A] to [B], [B] to [C] or [A] to [C]. While $F_B \approx F_C$ is the exact prediction of the theoretical model, the result that $F_A \approx F_B$ indicates that label II subjects do not view the sophistication of other label II subjects as significantly lower than their own, and therefore do not adjust their level of play in a measurable way. This result can therefore serve as a first step towards identifying subjects' beliefs over

[23] The only exception is at action 19, which is consistent with the well-known observation that stochastic dominance relations are often violated near the endpoints, even when the true distributions are stochastically-dominance ranked.

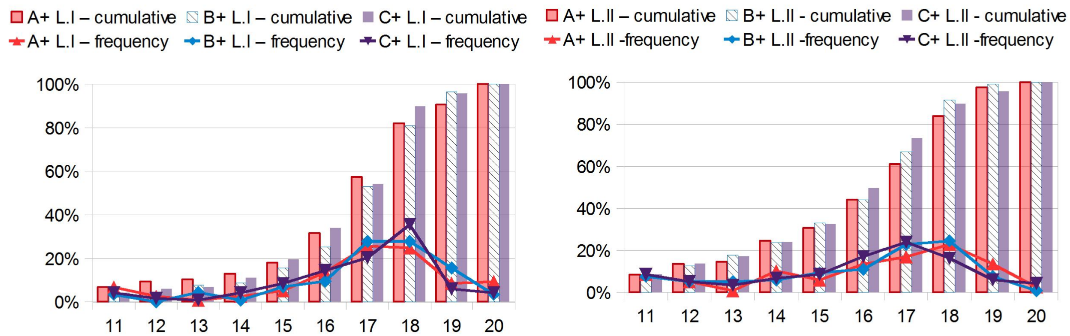


Figure 7: Treatments [A+], [B+] and [C+] for label I (left) and II (right).

their opponents.

We now compare the high payoff treatments [A+], [B+] and [C+], in which an individual whose action is exactly one below his opponent's receives an additional 80 tokens rather than 20. The model makes analogous predictions for these cases as it does for treatments [A], [B] and [C], namely that $F_{C+} \gtrsim F_{B+} \gtrsim F_{A+}$ for label I and $F_{A+} \gtrsim F_{B+} \approx F_{C+}$ for label II.

From Figure 7, no discernible pattern emerges either for label I or for label II, and we note that the (frequency) distributions are close to each other. None of the OLS estimates for the comparison of [A+] and [B+] or [A+] and [C+] are significant, for either label I or label II. The OLS estimates of [B+] and [C+] for label II are not significant either, although we note that they are significant from $[B+]$ to $[C+]$ ($< 5\%$) for label I in the other direction. Given the closeness of the distribution, this significance does not seem to be of first order.

The last results comparing treatments [A+], [B+] and [C+], viewed together with the results for treatments [A], [B] and [C], are indicative of label I subjects' beliefs. Specifically, these results suggest that label I subjects believe that the cost functions associated with label II subjects are higher than their own at low levels of k, but become closer to their own cost function at higher k's. In other words, label I subjects believe that, when sufficiently motivated, label II subjects are essentially the same as label I. An example of cost functions that satisfy this property is provided in Figure 2.a (p. 14). While the present analysis does not allow for precise identification of subjects' cost function, an extension of our approach could be used for this purpose.

4.3 Additional Observations

Our analysis of the results thus far has been from the viewpoint of testing our theory, which is the main goal of the experiment. We discuss next some findings that are not directly relevant to our model but that are useful for a broader understanding of individuals' behavior.

In Figure 9 of Appendix C.1, we report the realized payoffs for each action in the (modified) 11-20 game, computed using the empirical distributions observed in the various treatments for the two labels. Although distinct from the objectives of our theory, these realized payoffs allow for interesting observations. For instance, we find that the pure Nash equilibrium action in this game, 11, yields the lowest payoff in nearly all treatments, with optimal choices varying between 17 and 18. Therefore, a subject who has discovered the Nash equilibrium and plays accordingly would do worse (see Bosch-Domènech, Garcia-Montalvo, Nagel and Satorra (2002) for a discussion of this phenomenon). The rare occurrences of 11s and other low numbers in our data suggests that this kind of "curse of knowledge" is not particularly significant for the subjects of our experiment.

We also report the transition matrices for the low-high payoff comparisons, which serve to track individuals' behavior across different treatments (Figure 8 in Appendix C.1). Of particular interest in these matrices are the patterns of behavior concerning the choices of 20 and 11. Specifically, a plausible hypothesis is that 20 is chosen not only by level-0 players, but also by 'equilibrium players' who are not certain that the opponent would play 11 and have a strong degree of risk aversion (if the opponent plays 11, the payoffs from playing 11 and 20 is, respectively, 21 and 20). Observing large changes to and from 20 in the different treatments could be interpreted as evidence of this phenomenon.[24] This is not supported by the data, however, since transition matrices do not document a significant fraction of large changes to and from 20; the majority of these observations entail changes of two or three steps.[25]

5 Five 'Little Treasures' of Game Theory

In this section we show that our model sheds light on the well-known findings of Goeree and Holt (2001, henceforth GH). In this influential paper, GH conduct a series of experiments on initial responses in different games. For each of these games, GH contrast individuals' behavior in a baseline game, or 'treasure', with the behavior observed in a similar game, or 'contradiction', which differs only in the value of one parameter of the payoffs. GH show that classical equilibrium predictions often perform well in the treasure, but not in the contradiction. They also analyze whether existing models of strategic thinking can explain these findings, and conclude that they "do a fairly good job of organizing the qualitative patterns of conformity and deviation from the predictions of standard theory, but [that] there are obvious discrepancies" (GH, p. 1418). As they, and others since, note, it is important to have a model that explains these intuitive patterns of behavior. They further suggest that models of iterated introspection "offer some promise in explaining the qualitative features of deviations from Nash predictions" in their findings, but such an explanation has remained elusive. As we argue in this section, our model provides a unified explanation. We show that the qualitative predictions of the detail-free model are not only in line with the data in all the games within the domain of our theory (i.e., the five *static* games with *complete information*), but that our model also performs well quantitatively. Moreover, the central mechanism is in line with GH's observation that "the anomalous data patterns are related to the nature of incentives" (GH, p. 1416).

[24] We are grateful to one referee for this suggestion.

[25] We also note that the data from these matrices show that, at an individual level, roughly 80% of the observed changes are consistent with the theory.

Besides providing a unified explanation to GH's findings, this section illustrates how, with minimal added structure, our model can be used to perform 'structural' work on strategic reasoning and to obtain sharp predictions across games. These innovations of our approach open new directions of research. We first review GH's games and results, then present our explanation.

5.1 Little Treasures: Review

Matching Pennies. Consider the following game, with payoffs parameterized by $x > 40$:

	L	R
T	x,40	40,80
B	40,80	80,40

With $x = 80$, this is a standard Matching Pennies game. Nash Equilibrium predicts that both the row and the column players mix uniformly over their two actions. Since x does not affect the payoffs of the column player, in any Nash Equilibrium the distribution over the row player's actions should be uniform independent of x. While the equilibrium prediction is in line with the data observed when $x = 80$ (the 'treasure' treatment), when $x = 320$ or $x = 44$ (the 'contradiction' treatments), more than 95% of the row players choose the action with the relatively higher payoff: T when $x = 320$ and B when $x = 44$. Moreover, this behavior seems to have been anticipated by some of the column players, with roughly 80% percent of subjects playing the best response to the action played by most of the row players, which is R when $x = 320$ and L when $x = 44$.

Coordination Game with a Secure Outside Option. The following game, also parameterized by x, is a coordination game with one efficient and one inefficient equilibrium, which pay $(180, 180)$ and $(90, 90)$, respectively. The column player also has a secure option S which pays 40 independent of the row player's choice.

	L	H	S
L	90,90	0,0	x,40
H	0,0	180,180	0,40

Notice that action S is dominated by a uniform distribution over L and H. Hence, changing x has no effect on the set of equilibria. However, GH's experimental data show that behavior is strongly affected by x. In the treasure treatment ($x = 0$), a large majority of row and column players choose the efficient equilibrium action, and 80% of pairs coordinated on (H, H). In the contradiction treatment ($x = 400$), this percentage falls to 32%.

Traveler's Dilemma. In this version of Basu's (1994) well-known game, two players choose a number between 180 and 300 (inclusive). The reward they receive is equal to the lowest of their reports, but in addition the player who announces the higher number transfers a quantity

x to the other player.[26] This game is dominance solvable for any $x > 0$, and 180 is the only equilibrium strategy. GH observe that, when $x = 180$ (the 'treasure' treatment), roughly 80% choose numbers close to the Nash action, while when $x = 5$ (the 'contradiction' treatment), roughly 80% of subjects choose numbers close to the highest claim.

Minimum-Effort Coordination Game. Players in this game choose effort levels a_1, a_2 which can be any integer between 110 and 170. Payoffs are such that $u_i(a_1, a_2) = \min\{a_1, a_2\} - a_i \cdot x$, where x is equal to 0.1 in one treatment and 0.9 in the other. Independent of x, any common effort level is a Nash equilibrium. The efficient equilibrium is the one with high effort. While the pure-strategy Nash equilibria are unaffected by this change in payoffs, GH's experimental data show that agents exert lower effort when x is higher.

Kreps Game. The baseline and the modified games are described in the following table. The numbers in parenthesis represent the empirical distributions observed in the experiment:

Baseline:	*Left* (26)	*Middle* (8)	*Non Nash* (68)	*Right* (0)
Top (68)	$200, 50$	$0, 45$	$10, 30$	$20, -250$
Bottom (32)	$0, -250$	$10, -100$	$30, 30$	$50, 40$

Modified:	*Left* (24)	*Middle* (12)	*Non Nash* (64)	*Right* (0)
Top (84)	$500, 350$	$300, 345$	$310, 330$	$320, 50$
Bottom (16)	$300, 50$	$310, 200$	$330, 330$	$350, 340$

The modified game is obtained from the first simply by adding a constant of 300 to every payoff, which does not affect the equilibria. This game has two pure-strategy equilibria, $(Top, Left)$ and $(Bottom, Right)$, and one mixed-strategy equilibrium in which row randomizes between Top and $Bottom$ and column randomizes between $Left$ and $Middle$. Yet, a majority of column players choose the Non-Nash action. Also, in this case the change in payoffs has no effect on the column players, and only a small one on the row players.

5.2 Little Treasures: A Unified Explanation

The results of these experiments stand in sharp contrast with standard equilibrium concepts. Other heuristics, such as assuming that individuals play according to their 'maxmin' strategy, or based on risk or loss aversion, may explain the behavior observed in some games, but not in others. And yet, the observed behavior appears intuitive, and fundamentally linked to the nature of incentives. All five games share the feature that payoff parameter x changes between the treasure and contradiction treatments, without affecting the (pure-actions) best-reply functions. In the language of Section 2, this means that each treasure and its contradictions belong to the same cognitive equivalence class. Hence, we can use our model to understand the change of behavior by studying how varying the parameter x affects players' incentives to

[26] GH do not specify the rule in case of tie. We assume that there are no transfers in that case.

reason, holding the costs of reasoning constant.[27]

We show that, even in its detail-free version, our model improves our understanding of GH's findings for the Minimum-Effort game, the Traveler's Dilemma and Kreps' game. We then discuss how the model's predictions can be sharpened when some structure is added, and use the Matching Pennies and Coordination games to illustrate how our model can be used to perform a calibration exercise, which delivers sharp predictions across games.

5.2.1 Insights from the Detail-Free Model

We begin by considering the Minimum-Effort Coordination game. The predictions of our model for this game coincide with those of standard level-k models. That is because, for any given specification of the a^0, all levels $k \geq 1$ play the same action. Varying the cognitive bound thus has no behavioral implications. Following GH and appealing to the principle of insufficient reason, we set the anchor equal to the uniform distribution. Then, for any $k = 1, 2, ...$, we obtain $a_i^k = 164$ when $x = 0.1$ and $a_i^k = 116$ when $x = 0.9$. Our model therefore predicts that, independent of the shape of the cost and benefit functions (as long as $c(1) < v_i(1)$), players play 164 in the treasure and 116 in the contradiction. These results are close to the empirical findings, which are mainly concentrated near 170 and 110, respectively.

In the Traveler's Dilemma, GH's findings would be roughly consistent with level-1 reasoning if a^0 is uniform, since the best response would decrease from 290 to 180 as x increases from 5 to 180. They would not be consistent with a degenerate level-0 of $a_i^0 = 300$, however. We do not take a position as to whether or not the anchor is the uniform; our model is consistent with both. In particular, as x increases, Condition 2 implies that the value of reasoning increases. Hence, by Proposition 2, individuals depth of reasoning would be larger in the high reward treatment, and their chosen action would be lower. In this case, the observed change in behavior can be explained by the stronger 'incentives to reason' that the game provides when x is increased from 5 to 180. Another advantage of this explanation is that it does not require most players to be 'level-1'. As GH note, this is at odds with the classical findings of the literature on the distribution of the levels.

The remaining three games are very different in structure from the games considered thus far. Some of these appear sensitive to the specification of the anchor, and a theory of a^0 is beyond the scope of this paper. Our focus, however, is not on behavior in an isolated game, but on the comparative statics *across* games. For any a^0, our model entails restrictions across games. In the Kreps game, for instance, the implications of our model are the simplest of all. In particular, since the payoff differences are unchanged in the baseline and modified game, the value of reasoning remains the same (Condition 2). Hence, the model predicts that whatever we observe in the baseline game should not change for the modified game. This prediction is close to the observed behavior, especially for the column players.[28]

[27]We do not make cognitive equivalence assumptions other than for each treasure and its respective contradictions. For instance, the costs of reasoning for the Matching Pennies (and its contradictions) need not be the same as in the Traveler's Dilemma (and its contradiction).

[28]The distributions observed in the baseline game could, of course, be matched through appropriate choices

5.2.2 Extra Restrictions and Identification

Unlike the Kreps game, the value of reasoning for Matching Pennies and Coordination games may be different in the treasure treatment from the contradiction treatment. Analyzing these games therefore requires imposing extra structure on the model. In particular, we consider the following functional form for the value of reasoning:

$$v_i(k) = \max_{(a_i, a_{-i}) \in A} u_i(a_i, a_{-i}) - u_i\left(a_i^{k-1}, a_{-i}\right). \tag{8}$$

In words, the value of reasoning for player i, at each step, is equal to the maximum difference between the payoff that the player could get if he chose the optimal action a_i and the payoff he would receive given his current action a_i^{k-1}, out of all the possible opponent's actions. Effectively, individuals are optimistic over the gain in thinking more, or, alternatively, cautious about the validity of their current understanding.[29]

Imposing this particular functional form has no further implications for the Kreps and Minimum-Effort Coordination games. If instead we apply (8) to the Traveler's dilemma, and assume $a^0 = 300$, then the value of reasoning is equal to $v_i(k) = \max\{299 - a_i^{k-1}, 2x - 2\}$. Then if $x = 5$, $v_i(1) = 8$, and if $x = 180$, $v_i(k) = 358$ for $k < 120$ and 179 otherwise. With this value of reasoning, any cost function such that $c(5) > 8$ and $c(k) < 180$ for all k entails playing numbers higher than 295 when $x = 5$ and 180 when $x = 180$. This simple exercise shows how imposing a specific functional form for the value of reasoning enables a partial identification of the cost function. Identifying the cost of reasoning in different strategic settings is an important empirical question for future research.

A Calibration Exercise. We perform next a calibration exercise on the Matching Pennies and Coordination games, and show that with minimal added structure our model delivers sharp predictions across different games. Even though we allow for only one degree of freedom, calibrating a single parameter, our predictions fit the empirical findings closely, thereby providing strong support for our theory. Relative to the detail-free model, three kinds of assumptions are added to enable the identification necessary for this exercise:

(i) *On the functional forms:* We assume that the cost functions are strictly increasing and that the value of reasoning is specified as in (8).[30]

(ii) *On agents' beliefs:* We assume that there are two types of players, one (strictly) more sophisticated than the other (in the sense of Def. 2), respectively denoted by 'high' and 'low'. Let q_l denote the fraction of the low types. We maintain that q_l remains the same throughout all games and for both players. For identification purposes, we also assume that agents have

of other parameters of the model, but the focus of our analysis is in the comparative statics, according to which there should be no difference in behavior between the two treatments.

[29] An axiomatic foundation for this functional form is provided in Alaoui and Penta (2013).

[30] The results that follow do not rely on this specific functional form, which has the advantage of being restrictive in the degrees of freedom; our analysis performs at least as well if more degrees of freedom are allowed. For instance, our results would also hold if the value function were instead of the form $\sum_{a_{-i} \in A_{-i}} p(a_{-i})(u_i(BR_i(a_{-i}), a_{-i}) - u_i(a_i^{k-1}, a_{-i}))$, with positive probability $p(a_{-i})$ on each a_{-i}.

Data (GH)			*Calibration* $(q_l = 0.32)$		
$x = 80$	L (48)	R (52)	$x = 80$	L (50)	R (50)
T (48)			T (50)		
B (52)			B (50)		
$x = 320$	L (16)	R (84)	$x = 320$	L (16)	R (84)
T (96)			T (83)		
B (4)			B (17)		
$x = 44$	L (80)	R (20)	$x = 44$	L (84)	R (16)
T (2)			T (0)		
B (98)			B (100)		

Table 3: Matching Pennies

Data (GH)				*Calibration* $(q_l = 0.32)$			
$x = 0$	L (16)	H (84)	S	$x = 0$	L (17)	H (83)	S
L (4)				L (17)			
H (96)				H (83)			
$x = 400$	L	H	S	$x = 400$	L (33)	H (67)	S
L	(16)			L (49)	(16)		
H		(32)		H (51)		(34)	

Table 4: Coordination Game with a Secure Outside Option

correct beliefs over the distribution of types.

(iii) *On players' anchors:* While it is plausible that some anchors are more salient than others in some cases, we allow heterogeneity of anchors across individuals. We maintain throughout that, for all games, anchors are uniformly distributed in both populations and for all types. This neutrality ensures that our assumptions on the a^0 are neither post-hoc nor arbitrary.

The only parameter that we use in our calibration is the fraction of low types in the population, q_l. We calibrate this parameter using the Matching Pennies game with $x = 320$, and find $q_l = 0.32$. Maintaining throughout the parameter $q_l = 0.32$, we predict behavior for the other specifications of the Matching Pennies game (with $x = 80$ and $x = 44$), as well as for both the Coordination games, with $x = 0$ and $x = 400$. The results, shown in Tables 3 and 4, fit the data very well.[31]

Illustration of the Mechanism. For brevity, we leave the full details of the calibration to Appendix D. Here we illustrate the logic of the argument with a simple example. Consider

[31] Besides the data summarized in the matrices, GH also report that 64 percent of row players and 76 of the column players play H in the contradiction treatment of the coordination game ($x = 400$). These data however are inconsistent, probably due to a typographical error: if 76 percent of column play H and 32 percent of observations are (H, H), then cell (L, H) must receive a weight of 44. Since (L, L) is observed 16 percent of the times, it follows that at least 60 percent of row played L, which is inconsistent with 64 percent of Row playing H.

again the coordination game presented in Section 2, with $x > 0$:

	L	R
T	$x, 40$	$0, 0$
B	$0, 0$	$40, 40$

Using the maximum gain representation (8), the value of reasoning function for the row and column players, respectively indexed by $i = 1, 2$, is:

$$v_1(k) = \begin{cases} 40 & \text{if } a_1^{k-1} = T \\ x & \text{if } a_1^{k-1} = B \end{cases} \tag{9}$$
$$v_2(k) = 40.$$

Consistent with the assumptions above, players can be of two types, 'low' and 'high', with cost functions c_l and c_h, respectively, which are both increasing and such that $c_l \in C^-(c_h)$. We denote by q_l the fraction of low types, and assume that agents have correct beliefs about the distribution of types. In addition, assumption (iii) on the anchors entails that half of the players have 'Nash anchors' ((T, L) or (B, R)), and follow a degenerate path of reasoning; one quarter follow the path of reasoning $\{(T, R), (B, L), (T, R), ...\}$ and one quarter the path of reasoning $\{(B, L), (T, R), (B, L), ...\}$.

Now, suppose that $x = 40$. Then the value of reasoning is constant at 40 for both players. Given c_l, the low types' cognitive bound $\mathcal{K}(40, c_l) \equiv \hat{k}_{l_i}^{40}$ may be associated with any profile $a^{\hat{k}_{l_i}^{40}} \in \{(T, L), (T, R), (B, L), (B, R)\}$, depending on their a^0. Since the cognitive bound is binding for these types, their actions are distributed uniformly in both populations. The cognitive bound of the high types, $\hat{k}_{h_i}^{40} \equiv \mathcal{K}(40, c_h)$, is (weakly) higher than that of the low types: $\hat{k}_{h_i}^{40} \geq \hat{k}_{l_i}^{40}$. Hence, if $q_l > 1/2$, they best-respond to the low types, otherwise they play according to their own cognitive bound (cf. Example 4, p. 18). In either case, the high types' actions follow the distribution of the anchors. Hence, when $x = 40$, actions are uniformly distributed in both populations, as predicted by the unique mixed-strategy equilibrium. It is clear, however, that "*equilibrium play in this case is attained only by coincidence*" (cf. GH, p. 1407): with symmetric incentives (across players and across actions), behavior is completely driven by the anchors.

Now, suppose that x is increased, and consider the (non-Nash anchor) low types of population 1. For these players, the value of reasoning is no longer constant: $v_1(k+1)$ now alternates between 40 and $x > 40$, depending on whether the current action is $a_1^k = B$ or $a_1^k = T$, respectively. Players whose cognitive bound with $x = 40$ was such that $a^{\hat{k}_{l_i}^{40}} = (T, L)$ see no change in the value of the *next* step of reasoning. Their depth of reasoning therefore does not change either. For players who had stopped at $a^{\hat{k}_{l_i}^{40}} = (B, R)$ instead, the value of the next step is now x. Hence, they would perform one extra step if and only if $x > c_l(\hat{k}_{l_i}^{40} + 1)$.[32] If $q_l > 1/2$, this in

[32] These players perform one extra step, and no more, because $v(\hat{k}_{l_i}^{40} + 2) = 40 < c_l(\hat{k}_{l_i}^{40} + 1) < c_l(\hat{k}_{l_i}^{40} + 2)$.

turn implies that all the (non-Nash anchor) high types of population 2 play L, regardless of the anchor. Moreover, if q_l is not too low, then all the (non-Nash anchor) high types of population 1 play T. This threshold for q_l is decreasing in x, and approaches 0 as $x \to \infty$. Lastly, the low types of population 2 and players with Nash anchors behave as in the symmetric game.

Summarizing, for any pair of increasing functions $\left(c^l, c^h\right)$, if x is increased but below a certain threshold, then behavior does not change. Above that threshold, all the (non-Nash anchor) low types of population 1 play T. This change pins down the behavior of the high types in both populations which, for q_l in a suitable range, all coordinate on (T, L). Overall, greater coordination on (T, L) as x is increased above $x = 40$. Analogous results can also be proven if x is decreased below $x = 40$, in which case more players coordinate on (B, R))

We make the following observations. First, actions chosen for the low type players depend only on their own cognitive bound, while those of the high type players also depend on their beliefs over the low types' play. Second, for games that are symmetric in actions and payoffs, the distribution of anchors is identical to the distribution of play; if anchors are uniformly distributed, so are actions. Lastly, as payoffs are made asymmetric, incentives are distorted. For high enough payoff asymmetries between actions, the anchor itself is no longer relevant, provided that it is non-Nash: the behavior of the low types is driven by their incentives to reason, and becomes predictable. Depending on the parameter q_l, this in turn pins down the behavior of the high types. The calibration exercise essentially relies on this mechanism, which applies to all the variations of the Matching Pennies and Coordination games, and exploits the dependence of behavior on the distribution of types to identify the parameter q_l.[33] We also emphasize that the only property of the cost functions we used in this argument is that they are increasing. No further assumptions on the cost functions are needed to identify q_l.

6 Concluding Remarks

In this paper we have introduced a model of strategic thinking that endogenizes individuals' cognitive bounds as the result of a cost-benefit analysis. Our model also distinguishes between players' cognitive bounds and their beliefs about the opponent's bound. This differentiation raises several conceptual difficulties, such as reconciling higher order beliefs and bounded depth of reasoning, especially when the latter is endogenous. Our general framework accounts for the interactions between depth of reasoning, incentives and general hierarchies of beliefs. We have also proposed a simplified model, which provides a tractable framework to guide the design of an empirical test of these complex interactions.

From a theoretical viewpoint, we extend the general level-k approach of taking reasoning in games to be procedural and possibly constrained. By making explicit an appealing feature of level-k models that play follows from a reasoning procedure, our framework serves to attain

The first inequality follows from the definition of $\hat{k}^{40}_{l_i}$, and the second from c_l being increasing.

[33] For instance, in this example, for x sufficiently large we obtain the following: $3/4$ of population 1 plays T and $1/4$ plays B irrespective of q_l; $1/2$ of population 2 plays L and the other half plays R if $q_l \leq 1/2$; and $\left(\frac{3}{4} - \frac{q_l}{4}\right)$ plays L and the others play R if instead $q_l > 1/2$.

a deeper understanding of the underlying mechanisms of that approach. Our framework also solves apparent conceptual difficulties of the level-k approach, such as the possibility that individuals reason about opponents they regard as more sophisticated.

In addition to testing the 'detail free' model, our experiment plays a more general role. It reveals that individuals change their actions as their incentives and beliefs about the opponents are varied. Players change their behavior in a systematic way, not captured by existing models of strategic reasoning. Taken together, these findings suggest that the fundamental reasoning process behind individual choices follows clear patterns. Moreover, observed behavior is not only a function of individuals' strategic sophistication, but also of incentives and beliefs. Thus, caution should be exercised in interpreting level of play as purely revealing of cognitive ability, as an endogeneity problem is present.[34] These results demonstrate the need for a richer theory of procedural rationality under non-equilibrium play, for which our model serves as a natural and tractable candidate.

After having verified the empirical validity of the broader cost-benefit approach, we have imposed structure to the model to analyze Goeree and Holt's (2001) influential 'little treasures' experiments. The added structure uses the representation of the cost-benefit analysis axiomatically derived in Alaoui and Penta (2013). Our theoretical predictions for these games are highly consistent with the empirical findings, thereby providing further evidence in support of our theory, and an external validation of our approach. Moreover, since Goeree and Holt's experiments include games that have a very different structure from the games used in our experiment, our calibration exercise also shows that our theory is applicable to a wide range of games.

In closing, we note that by relating individuals' cognitive bounds to their incentives in the game, our theory establishes a link between level-k reasoning and the conventional domain of economics, centered around tradeoffs and incentives. We see this as a desirable feature from a methodological viewpoint, which can further favor the integration of theories of initial responses within the core of economics. Conversely, the application of classical economic concepts to a model of reasoning opens new directions of research from both a theoretical and an empirical viewpoint. Two related avenues of research, outlined in Section 5, include a rigorous identification of the properties of cost functions in different games and testing predictions of the direction and magnitudes of changes in behavior across other strategic settings. These last questions are outside the scope of this paper, and remain open for future research.

[34] In a different setting, it is a well known theme in the Economics of Education literature that incentives may affect standard measures of cognitive abilities. For a recent survey of the vast literature that combines classical economics notions with measurement of cognitive abilities and psychological traits to address the endogeneity problems stemming from the role of incentives, see, for instance, Almlund, Duckworth, Heckman and Kautz (2011).

References

1. Agranov, Marina, Andrew Caplin and Chloe Tergiman. 2012. "Naive Play and the Process of Choice in Guessing Games" *mimeo.*

2. Agranov, Marina, Elizabeth Potamites, Andrew Schotter and Chloe Tergiman. 2012. "Beliefs and Endogenous Cognitive Levels: An Experimental Study" *Games and Economic Behavior*, 75(2): 449-63.

3. Alaoui, Larbi and Antonio Penta. 2013. "Strategic Thinking and the Value of Reasoning," *mimeo.*

4. Almlund, Mathilde, Angela Lee Duckworth, James Heckman and Tim Kautz. 2011. "Personality Psychology and Economics", *Handbook of the Economics of Education*, Vol. 4.

5. Arad, Ayala and Ariel Rubinstein. 2012. "The 11-20 Money Request Game: A Level-k Reasoning Study", *American Economic Review*, 102(7): 3561-3573.

6. Basu, Kaushik. 1991. "The Traveler's Dilemma: Paradoxes of Rationality in Game Theory." *American Economic Review Papers and Proceedings*, 84(2): 391-395.

7. Bhatt, Meghana A., and Colin F. Camerer. 2005. "Self-referential Thinking and Equilibrium as States of Mind in Games: fMRI Evidence." *Games and Economic Behavior*, 52(2): 424-459.

8. Bhatt, Meghana A., Terry Lohrenz, Colin F. Camerer, and P. Read Montague. 2010. "Neural Signatures of Strategic Types in a Two-Person Bargaining Game." *Proceedings of the National Academy of Sciences*, 107(46): 19720-19725.

9. Bosch-Domènech, Antoni, Jose García-Montalvo, Rosemarie Nagel, and Albert Satorra. 2002. "One, Two, (Three), Infinity...: Newspaper and Lab Beauty-Contest Experiments." *American Economic Review*, 92(5), 1687-1701.

10. Camerer, Colin F., Teck-Hua Ho, and Juin Kuan Chong. 2004. "A Cognitive Hierarchy Model of Games." *Quarterly Journal of Economics*, 119(3): 861-898.

11. Capra, C. Monica, Jacob K. Goeree, Rosario Gomez, and Charles A. Holt. 1999. "Anomalous Behavior in a Traveler's Dilemma?" *American Economic Review*, 89(3): 678-690.

12. Choi, Syungjoo. 2012. "A Cognitive Hierarchy Model of Learning in Networks." *Review of Economic Design*, 16: 215-250.

13. Coricelli, Giorgio, and Rosemarie Nagel. 2009. "Neural Correlates of Depth of Strategic Reasoning in Medial Prefrontal Cortex." *Proceedings of the National Academy of Sciences*, 106(23): 9163-9168.

14. Costa-Gomes, Miguel A., and Vincent P. Crawford. 2006. "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study." *American Economic Review*, 96(5): 1737-1768.

15. Costa-Gomes, Miguel A., Vincent P. Crawford, and Bruno Broseta. 2001. "Cognition and Behavior in Normal-Form Games: An Experimental Study." *Econometrica*, 69(5): 1193-1235.

16. Crawford, Vincent P. 2003. "Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions." *American Economic Review,* 93(1): 133-149.

17. Crawford, Vincent P., Miguel A. Costa-Gomes, and Nagore Iriberri. 2013. "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications." *Journal of Economic Literature*, 51.

18. Crawford, Vincent P., and Nagore Iriberri. 2007. "Level-k Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?" *Econometrica*, 75(6): 1721-1770.

19. Davidson, Russell, and Jean-Yves Duclos. 2000. "Statistical inference for stochastic dominance and for the measurement of poverty and inequality." *Econometrica*, 68(6): 1435-1464.

20. Fehr, Dietmar., and Steffen Huck. 2013. "Who knows it is a game? On rule understanding, strategic awareness and cognitive ability." *mimeo.*

21. Gabaix, Xavier. 2012. "Game Theory with Sparsity-Based Bounded Rationality." *mimeo.*

22. Gill, David., and Victoria Prowse. 2013. "Cognitive ability and learning to play equilibrium: A level-k analysis." *mimeo.*

23. Georganas, Sotiris., Paul, J. Healy and Roberto A. Weber. 2012. "Cognitive ability and learning to play equilibrium: A level-k analysis." *mimeo.*

24. Goeree, Jacob K., and Charles A. Holt. 2001. "Ten Little Treasures of Game Theory and Ten Intuitive Contradictions." *American Economic Review*, 91(5): 1402-1422.

25. Goeree, Jacob K., and Charles A. Holt. 2004. "A Model of Noisy Introspection." *Games and Economic Behavior*, 46(2): 365-382.

26. Grosskopf, Brit, and Rosemarie Nagel. 2008. "The Two-Person beauty contest." *Games and Economic Behavior* , 62(1): 93-99.

27. Kets, Willemien. 2012. "Bounded Reasoning and Higher-Order Uncertainty." *mimeo*, Kellogg School of Management, Northwestern University.

28. Kneeland, Terri. 2012. "Coordination under Limited Depth of Reasoning." *mimeo*, University of British Columbia.

29. Ho, Teck-Hua, Colin Camerer, and Keith Weigelt. 1998. "Iterated Dominance and Iterated Best Response in Experimental 'p-Beauty Contests'." *American Economic Review*, 88(4): 947-969.

30. Nagel, Rosemarie. 1995. "Unraveling in Guessing Games: An Experimental Study." *American Economic Review*, 85(5): 1313-1326.

31. Palacios-Huerta, Ignacio and Oscar Volij. 2009. "Field Centipedes." *American Economic Review*, 99(4): 1619-1635.

32. Penta, Antonio. 2012. "Higher Order Uncertainty and Information: Static and Dynamic Games," *Econometrica*, Vol. 80(2), 631-660

33. Penta, Antonio. 2013. "On the Structure of Rationalizability on Arbitrary Spaces of Uncertainty", *Theoretical Economics*, 8, 405–430.

34. Stahl, Dale O., and Paul R. Wilson. 1994. "Experimental Evidence on Players' Models of Other Players." *Journal of Economic Behavior and Organization*, 25(3): 309-327.

35. Stahl, Dale O., and Paul R. Wilson. 1995. "On Players' Models of Other Players: Theory and Experimental Evidence." *Games and Economic Behavior*, 10(1): 218-254.

36. Schipper, Burkhard C.. 2014. "Awareness." *mimeo*, UC-Davis.

37. Strzalecki, Tomasz. 2010. "Depth of Reasoning and Higher-Order Beliefs." Harvard Institute of Economic Research Discussion Paper Number 2184.

38. Weinstein, Jonathan L. and Muhamet Yildiz. 2007. "A Structure Theorem for Rationalizability With Application to Robust Predictions of Refinements," *Econometrica*, 75, 365–400.

39. Weinstein, Jonathan L. and Muhamet Yildiz. 2013. "Robust Predictions in Infinite-horizon Games–An Unrefinable Folk Theorem," *Review of Economic Studies,* forthcoming.

Appendix

A Logistics of the Experiment

The experiment was conducted at the Laboratori d'Economia Experimental (LEEX) at Universitat Pompeu Fabra (UPF), Barcelona. Subjects were students of UPF, recruited using the LEEX system. No subject took part in more than one session. Subjects were paid 3 euros for showing up (students coming from a campus that was farther away received 4 euros instead). Subjects' earnings ranged from 10 to 40 euros, with an average of 15.8.

Each subject went through a sequence of 18 games. Payoffs are expressed in 'tokens', each worth 5 cents. Subjects were paid randomly, once every six iterations. The order of treatments is randomized (see below). Finally, subjects only observed their own overall earnings at the end, and received no information concerning their opponents' results.

Our subjects were divided in 6 sessions of 20 subjects, for a total of 120 subjects. Three sessions were based on the exogenous classification, and each contained 10 students from the field of humanities (humanities, human resources, and translation), and 10 from math and sciences (math, computer science, electrical engineering, biology and economics). Three sessions were based on the endogenous classification, and students were labeled based on their performance on a test of our design. (See Appendix B). In these sessions, half students were labeled as 'high' and half as 'low'.

A.1 Instructions of the Experiment

We describe next the instructions as worded for a student from math and sciences. The instructions for students from humanities would be obtained replacing these labels everywhere. Similarly, labels high and low would be used for the endogenous classification.

A.1.1 Baseline Game and Treatments [A], [B] and [C]

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.

Example:

-If you say 17 and your opponent says 19, then you receive 17 and he receives 19.

-If you say 12 and your opponent says 13, then your receive 32 and he receives 13.

-If you say 16 and you opponent says 16, then you receive 26 and he receives 26.

Treatments [A] and [B]:

Your opponent is:

- a student from maths and sciences (treatment [A]) / humanities (treatment [B])

- he is given the same rules as you.

Treatment [C]:

In this case, the number you play against is chosen by:

- a student from humanities facing another student from humanities. In other words, two students from humanities play against each other. You play against the number that one of them has picked.

A.1.2 Changing Payoffs: Treatments [A+], [B+] and [C+]

You are now playing a high payoff game.

- if you give the same number as your opponent, you receive an extra 10 .

Example:

-If you say 17 and your opponent says 19, then you receive 17 and he receives 19.

-If you say 12 and your opponent says 13, then you receive 92 and he receives 13.

-If you say 16 and you opponent says 16, then you receive 26 and he receives 26.

Treatments [A+] and [B+]

Your opponent is:

- a student from maths and sciences playing the high payoff game (treatment [A+]) / humanities (treatment [B+])

- he is given the same rules as you.

Treatment [C+]

In this case, the number you play against is chosen by:

- a student from humanities playing the high payoff game with another student from humanities. In other words, two students from humanities play the high payoff game with each other (extra 10 if they tie, 80 if exactly one less than opponent). You play against the number that one of them has picked.

A.1.3 Treatments [D], [E] and [F]

(The results of these treatments are contained in appendix E in the Supplementary materials.) Before playing treatments [D], [E] and [F], the subjects were given the following 'tutorial':

Game Theory Tutorial: According to game theory, if the players are infinitely rational, then the game should be played in the following way. Both players should say 11.

Explanation: Suppose the two players are named Ana and Beatriz. If Ana thinks Beatriz plays 20, then Ana would play 19. But then Beatriz knows that Ana would play 19, so she would play 18. Ana realizes this, and so she would play 17.... they both follow this reasoning until both would play 11. Notice that if Beatriz says 11, then the best thing for Ana is to also say 11.

Treatment [D]

Your opponent is:

- a student who has also been given the game theory tutorial.

Treatment [E]

Your opponent is:

- a student from maths and sciences.

- he has not been given the game theory tutorial.

Treatment [F]

Your opponent is:

- a student from humanities.

- he has not been given the game theory tutorial.

A.1.4 Treatments [K] and [L]

(The results of these treatments are contained in appendix E in the Supplementary materials.)

Treatment [K]

In this case, the number you play against is chosen by:

- a student from maths and sciences playing the low payoff game with another student from maths and sciences. In other words, two students from maths and sciences play the low payoff game with each other (extra 10 if they tie, 20 if exactly one less than opponent). You play against the number that one of them has picked.

Treatment [L]

In this case, the number you play against is chosen by:

- a student from humanities playing the low payoff game with another student from humanities. In other words, two students from humanities play the low payoff game with each other (extra 10 if they tie, 20 if exactly one less than opponent). You play against the number that one of them has picked.

A.2 Sequences

Our 6 groups (3 for the endogenous and 3 for the exogenous classification) went through four different sequences of treatments. Two of the groups in the exogenous treatment followed Sequence 1, and one followed Sequence 2. The three groups of the endogenous classification each took a different sequence: respectively sequence 1, 3 and 4. All the sequences contain our main treatments, [A], [B], [C] ,[A+], [B+], [C+]. In addition, sequences 2 and 4 contain the [K] and [L] treatments, whereas sequences 1 and 3 conclude with the additional treatments [D], [E] and [F]. The order of the main treatments is different in each sequence, both in terms of changing the beliefs and the payoffs.

- **Sequence 1:** $A, B, C, B, A, C, A^+, B^+, C^+, B^+, A^+, C^+, D, E, F, D, E, F$

- **Sequence 2:** $A, B, B, A, C, C, K, L, K, L, K, L, A^+, B^+, B^+, A^+, C^+, C^+$

- **Sequence 3:** $A^+, B^+, C^+, B^+, A^+, C^+, A, B, C, B, A, C, D, E, F, D, E, F$

- **Sequence 4:** $B, A, C, B, A, C, K, L, K, L, K, L, B^+, A^+, C^+, B^+, A^+, C^+$

B The Test for the Endogenous Classification

The cognitive test takes roughly thirty minutes to complete, and consists of three questions. In the first, subjects are asked to play a variation of the board game Mastermind. In the second question, the subjects are given a typical centipede game of seven rounds, and are asked what an infinitely sophisticated and rational agent would do. In the third game, the subjects are given a lesser known 'pirates game', which is a four player game that can be solved by backward induction. Subjects are asked what the outcome of this game would be, if players were 'infinitely sophisticated and rational'. Each question was given a score, and then a weighted average was taken. Subjects whose score was higher (lower) than the median score were labeled as 'high' ('low'). We report next the instructions of the test, as administered to the students (see the online appendix for the original version in Spanish).

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

Question 1:

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 second per round: 30 seconds to introduce the numbers and 60 seconds to view the results.

Question 2:

Consider the following game.Two people, Antonio and Beatriz, are moving sequentially. The game starts with 1 euro on the table. There 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.

Question 3:

Four pirates (Antonio, Beatriz, Carla and David) have obtained 10 gold doblónes and have to divide up the loot. Antonio proposes a distribution of the loot. All pirates vote on the proposal. If half the crew or more agree, the loot is divided as proposed by Antonio.

If Antonio fails to obtain support of at least half his crew (including himself), then he will be killed. The pirates start over again with Beatriz as the proposer. If she gets half the crew (including herself) to agree, then the loot is divided as proposed. If not, then she is killed, and Carla then makes the proposal. Finally, if her proposal is not agreed on by half the people left, including herself, then she is killed, and David takes everything.

In other words:

Antonio needs 2 people (including himself) to agree on his proposal, and if not he is killed.

If Antonio is killed, Beatriz needs 2 people (including herself) to agree on her proposal, if not she is killed.

If Beatriz is killed, Carla needs 1 person to agree (including herself) to agree on her proposal, and if not she is killed.

If Carla is killed, David takes everything.

The pirates are infinitely sophisticated and rational, and they each want to get as much money as possible. What is the maximum number of coins Antonio can keep without being killed?

Notice that *the proposer* can also vote, and that exactly half the votes is enough for the proposal to pass.

You have 8 minutes in total for this question.

Scoring. In the *mastermind* question, subjects were given 100 points if correct, otherwise they received 15 points for each correct answer in the correct place and 5 for each correct answer in the wrong place in their last answer. In the *centipede* game, subjects were given 100 points if they answered that the game would end at round 1, otherwise points were equal to $\min\{0, (6 - round) \cdot 15\}$. In the *pirates* game, subjects obtain 100 if they answer 100, 60 if they answer 10, $\min\{0, (80 - x) \cdot 10\}$. The overall score was given by the average of the three.

C Statistical Tests and Regressions

	Relevant dummy	Classification dummy	Constant	Number of obs.
From A to A+ Label I	-0.50*** (0.17)	0.22 (0.52)	17.21	235
From B to B+ Label I	-0.62*** (0.19)	0.36 (0.38)	17.50	233
From C to C+ Label I	-1.15*** (0.18)	0.34 (0.38)	17.76	236
From A to A+ Label II	-0.64** (0.27)	-0.10 (0.45)	16.91	236
From B to B+ Label II	-0.74*** (0.25)	0.38 (0.47)	16.57	236
From C to C+ Label II	-0.97*** (0.25)	-0.07 (0.45)	16.97	234
From B to A Label I	-0.36* (0.20)	0.34 (0.43)	17.50	236
From B to C Label I	0.25 (0.25)	0.44 (0.37)	17.46	236
From A to C Label I	0.62*** (0.18)	0.36 (0.42)	17.13	236
From A to B Label II	-0.09 (0.26)	0.50 (0.48)	16.6	236
From B to C Label II	0.16 (0.25)	0.36 (0.46)	16.59	235
From A to C Label II	0.07 (0.27)	0.30 (0.44)	16.71	235
From B+ to A+ Label I	-0.26 (0.17)	0.25 (0.46)	16.96	232
From B+ to C+ Label I	-0.29** (0.13)	0.27 (0.45)	16.93	233
From A+ to C+ Label I	0.03 (0.16)	0.21 (0.48)	16.71	236
From A+ to B+ Label II	-0.18 (0.21)	-0.22 (0.52)	16.33	236
From B+ to C+ Label II	-0.07 (0.25)	-0.04 (0.50)	16.05	235
From A+ to C+ Label II	-0.25 (0.21)	-0.47 (0.51)	16.46	235

Table 5: Regressions for Labels I and II. * indicates $<$ 10% significance, ** indicates $<$ 5% significance and *** indicates $<$ 1% significance. Standard errors in parenthesis.

	Two-Sample Kolmogorov-Smirnov D-stat (exact p-value)	Wilcoxon-signed-rank p-values
A vs A+ Label *I*	0.18 (0.040) **	0.0002 ***
B vs B+ Label *I*	0.24 (0.002) ***	0.0001 ***
C vs C+ Label *I*	0.37 (0.000) ***	0.0000 ***
A vs A+ Label *II*	0.14 (0.128)	0.0020 ***
B vs B+ Label *II*	0.17 (0.048) **	0.0013 ***
C vs C+ Label *II*	0.25 (0.001) ***	0.0001 ***

Table 6: Equality of Distributions Tests: Changing Payoffs. * indicates < 10% significance, ** indicates < 5% significance and *** indicates < 1% significance.

	Two-Sample Kolmogorov-Smirnov D-stat (exact p-value)	Wilcoxon-signed-rank p-values
A vs B Label *I*	0.11 (0.463)	0.37
B vs C Label *I*	0.05 (0.998)	0.37
A vs C Label *I*	0.14 (0.163)	0.02 **
A+ vs B+ Label *I*	0.06 (0.952)	0.08 *
B+ vs C+ Label *I*	0.08 (0.738)	0.03 **
A+ vs C+ Label *I*	0.07 (0.869)	0.82
A vs B Label *II*	0.05 (0.986)	0.88
B vs C Label *II*	0.06 (0.957)	0.73
A vs C Label *II*	0.04 (1.000)	0.79
A+ vs B+ Label *II*	0.07 (0.793)	0.17
B+ vs C+ Label *II*	0.06 (0.938)	0.23
A+ vs C+ Label *II*	0.12 (0.277)	0.08 *

Table 7: Equality of Distributions Tests: Changing Opponents. * indicates < 10% significance, ** indicates < 5% significance and *** indicates < 1% significance.

	From A to A+ Label I (t-statistics)	From B to B+ Label I (t-statistics)	From C to C+ Label I (t-statistics)
11	$-.0441$	$-.0967$	$-.1050$
12	$-.1344$	$-.0039$	$-.1172$
13	$-.0956$	$-.0547$	$-.1401$
14	$-.0268$	$-.0298$	$-.1394$
15	$-.0449$	$-.0732$	$-.2549$
16	$-.1360$	$-.1227$	$-.3846$
17	$-.2763$	$-.4009$	$-.5312$
18	$-.3625$	$-.4593$	$-.6789$
19	$-.0886$	$-.1340$	$-.2222$
20			

Table 8: DD test: Changing Payoffs, Label I

	From A to A+ Label II (t-statistics)	From B to B+ Label II (t-statistics)	From C to C+ Label II (t-statistics)
11	$-.0836$	$-.0768$	$-.2209$
12	$-.0871$	$-.0635$	$-.1501$
13	$-.0412$	$-.0393$	$-.0893$
14	$-.1273$	$-.0553$	$-.1858$
15	$-.0652$	$-.1313$	$-.1650$
16	$-.1338$	$-.2267$	$-.1989$
17	$-.2245$	$-.2812$	$-.4879$
18	$-.2520$	$-.3309$	$-.3865$
19	$-.3090$	$-.3249$	$-.2185$
20			

Table 9: DD test: Changing payoffs, Label II

C.1 Supplementary Descriptive Statistics

A vs A+	11	12	13	14	15	16	17	18	19	20
11	8				1	1		2	2	
12	2							2	1	
13		2				2			1	
14		1		2	1	3	1	1	2	
15		1		6	4		3	2		
16	1	2		3	2	7	3	3	1	
17	2			3	1	7	9	7	3	
18	2	2	2	1	2	7	16	13	5	2
19	3				1	4	11	22	6	2
20		1			1	1	7	4	5	10

changes in the direction consistent with the theory: 78%

B vs B+	11	12	13	14	15	16	17	18	19	20
11	2	2	1	1	1		1			
12	1		2				1	2	2	
13	2			1	2	2	1	1	1	
14	1	1		2	1		1	2		
15			1	1	4	3	2	1		
16		1	1		1	4	5	1	1	
17	3	1	2	2	3	4	8	8	3	1
18	1	1	3	1	6	6	20	18	2	
19	2		1		1		14	20	16	3
20	1					5	6	8	2	1

changes in the direction consistent with the theory: 76%

C vs C+	11	12	13	14	15	16	17	18	19	20
11	2					1	1		1	
12		1		1	1		1	3		
13	1				1	1	2	1		
14		1		3	1			1		
15	2		1		5	2	3	1		
16	1	1	2	2	4	10	4	1		
17	2	3		2	3	7	9			
18	1	1	1	4	3	5	13	22	4	3
19	4	1	1	1	1	8	14	24	4	3
20	2				1	3	5	8	5	4

changes in the direction consistent with the theory: 84%

Figure 8: Transition Matrices

payoff functions

	A L.I	B L.I	C L.I		A+ L.I	B+ L.I	C+ L.I
11	**11.59**	**12.53**	**12.44**		13.74	**15.83**	15.92
12	12.34	13.69	12.93		**12.94**	16.58	**13.19**
13	14.02	14.61	14.27		15.14	18.25	21.22
14	15.27	15.53	16.37		18.36	22.05	19.76
15	17.29	16.53	17.80		26.45	24.75	26.44
16	20.32	20.07	18.97		37.88	35.41	30.92
17	23.10	**24.03**	22.59		**39.39**	**38.95**	**37.00**
18	**24.95**	23.17	**23.85**		27.32	26.56	31.14
19	23.75	22.64	23.32		27.38	20.44	22.39
20	21.19	21.19	21.27		20.94	20.08	20.25

	A L.II	B L.II	C L.II		A+ L.II	B+ L.II	C+ L.II
11	**12.44**	**11.51**	**11.59**		15.92	**11.35**	13.74
12	12.93	12.68	12.34		**13.19**	15.48	**12.94**
13	14.27	13.59	14.02		21.22	14.13	15.14
14	16.37	15.19	15.27		19.76	19.65	18.36
15	17.80	16.86	17.29		26.44	23.35	26.45
16	18.97	19.05	20.32		30.92	39.22	37.88
17	22.59	23.27	23.10		**37.00**	**42.04**	**39.39**
18	**23.85**	**27.66**	**24.95**		31.14	33.30	27.32
19	23.32	24.08	23.75		22.39	23.35	27.38
20	21.27	20.76	21.19		20.25	20.35	20.94

SUMMARY

	A L.I	B L.I	C L.I	A+ L.I	B+ L.I	C+ L.I
average	17.32	17.69	17.94	16.81	17.05	16.79
modal	19	19	19	17	17-18	18
best	18	17	18	17	17	17
max expec	24.95	24.03	23.85	39.39	38.95	37.00
	A L.II	B L.II	C L.II	A+ L.II	B+ L.II	C+ L.II
average	16.86	16.77	16.93	16.22	16.03	15.96
modal	18	18	18	18	18	17
best	18	18	18	17	17	17
max expec	23.85	27.66	24.95	37.00	42.04	39.39

Figure 9: Realized Payoffs and Summary Statistics

D Calibration

D.1 Matching Pennies

Given the functional form in (8), the value of reasoning is:

$$v_1(k) = \begin{cases} 40 & \text{if } a_1^{k-1} = T \\ x - 40 & \text{if } a_1^{k-1} = B \end{cases} \tag{10}$$

$$v_2(k) = 40.$$

The paths of reasoning have a periodicity of 4. For instance, for $a^0 = (B, L)$, the path is $(B, L), (T, L), (T, R), (B, R), (B, L), \ldots$. The cases for the other possible three anchors are obtained similarly.

Fix cost functions $\left(c^l, c^h\right)$, and let $x = 80$. Then, $v_i(k) = 40$ for all i and k. Let v^{80}denote such function. Define $k^h = \mathcal{K}\left(c^h, v^{80}\right)$, $k^l = \mathcal{K}\left(c^l, v^{80}\right)$ and $\Delta k := k^h - k^l$. Under the maintained assumptions, $\Delta k \geq 1$. Given the symmetry of the incentives and the uniform distribution of the anchors, with $x = 80$ we obtain a uniform distribution over actions in both populations, independent of the value of q_l.

We now consider the case $x = 320$, assuming that it is sufficiently high to affect behavior. Then, by the same argument of Section 5.2.2, the path of reasoning for all agents in population 1 stops at a point in which they consider action T. Also notice that, with this payoffs, agents of population 1prefer action T as soon as they attach probability at least $1/8$ to the opponents playing L. The cut-off probability for population 2 remains at $1/2$.

1. Suppose that $q_l < 1/8$. When $q_l < 1/8$, the behavior of the low types is not enough to pin down the behavior of the high types in either population. It follows that all types will play at their cognitive bound. In particular, this implies that the actions of both types in population 2 are uniformly distributed. The prediction that population 2 plays uniformly with $x = 320$ is inconsistent with the data. We thus rule out the case $q_l < 1/8$.

2. Suppose that $q_l > 1/2$. Since, independent of the anchor, all low types of population 1 switch to T, then all high types of population 2 play R if $q_l > 1/2$. The low types of population 2instead are uniformly distributed, following their own cognitive bound. It follows that a fraction $\frac{q_l}{2} + (1 - q_l)$ of population 2plays R. Given the uniformity assumption on the anchors, half of the high types in population 1 believe that the low types of population 2 play L, and since $q_l > 1/8$, they play T. The remaining half of high types of population 1 believe the low types of population 2 play R, hence they play B. The resulting distribution is the following:

   for $q_l > 1/2$:

$x = 320$	$L\ \left(\frac{q_l}{2}\right)$	$R\ \left(1 - \frac{q_l}{2}\right)$
$T\ \left(\frac{1}{2} + \frac{q_l}{2}\right)$		
$B\ \left(\frac{1}{2} - \frac{q_l}{2}\right)$		

3. Suppose that $q_l \in (1/8, 1/2)$. Given c_l, depending on what the anchor is, we may have the following cases:

   (a) $a^{k^l} = (B, L)$ or $a^{k^l} = (T, L)$. In this case, which applies to half of the population, the low types of population 2 play L. Since $q_l > 1/8$, this is enough to convince the high types of population 1 to play T. The low types of population 1 play T, because their increased incentives moved their cognitive bound to T.

      i. If $\Delta k \geq 2$, then the high types of population 2 understand everything thus far, hence play R.

      ii. If $\Delta k = 1$, then the high types of population 2 are not sufficiently 'deep' to understand the choice of the high types of population 1 (which are best responding to the low types of population 2). They thus play at their bound. Whether this is L or R depends on the anchor being (B, L) or (B, R), which is uniformly distributed.

(b) $a^{k^l} = (B, R)$. In this case, which applies to a quarter of the population, the low types of population 2 play R and the low types of population 1 play T(because of the increased incentives, they stop at step $k^l + 1$). Since $q_l \in (1/8, 1/2)$, these are not enough to pin down the behavior of the high types in either population. The high types of both populations therefore play at their bound, which is (B, L) if $\Delta k = 1 \,(\mathrm{mod}\, 4)$, (T, L) if $\Delta k = 2 \,(\mathrm{mod}\, 4)$, (T, R) if $\Delta(k) = 3 \,(\mathrm{mod}\, 4)$ and (B, R) if $\Delta k = 4 \,(\mathrm{mod}\, 4)$.

(c) $a^{k^l} = (T, R)$. In this case, which applies to a quarter of the population, the low types of population 2 play R and the low types of population 1 play T. Since $q_l \in (1/8, 1/2)$, these are not enough to pin down the behavior of the high types in either population. The high types of both populations therefore play at their bound, which is (B, R) if $\Delta k = 1 \,(\mathrm{mod}\, 4)$, (B, L) if $\Delta k = 2 \,(\mathrm{mod}\, 4)$, (T, L) if $\Delta(k) = 3 \,(\mathrm{mod}\, 4)$ and (T, R) if $\Delta k = 4 \,(\mathrm{mod}\, 4)$.

Aggregating cases (a),(b) and (c), we have the following possibilities for $q_l \in (1/8, 1/2)$:

for $q_l \in (1/8, 1/2)$, $\Delta k = 1$

$x = 320$	$L\ \left(\frac{1}{2}\right)$	$R\ \left(\frac{1}{2}\right)$
$T\ \left(\frac{1}{2} + \frac{q_l}{2}\right)$		
$B\ \left(\frac{1}{2} - \frac{q_l}{2}\right)$		

for $q_l \in (1/8, 1/2)$, $\Delta k \geq 2$:

$\Delta k = 1 \,(\mathrm{mod}\, 4)$:

$x = 320$	$L\ \left(\frac{1}{4} + \frac{q_l}{4}\right)$	$R\ \left(\frac{3}{4} - \frac{q_l}{4}\right)$
$T\ \left(\frac{1}{2} + \frac{q_l}{2}\right)$		
$B\ \left(\frac{1}{2} - \frac{q_l}{2}\right)$		

$\Delta k = 2 \,(\mathrm{mod}\, 4)$:

$x = 320$	$L\ (1/2)$	$R\ (1/2)$
$T\ \left(\frac{3}{4} + \frac{q_l}{4}\right)$		
$B\ \left(\frac{1}{4} - \frac{q_l}{4}\right)$		

$\Delta k = 3 \,(\mathrm{mod}\, 4)$:

$x = 320$	$L\ \left(\frac{1}{4} + \frac{q_l}{4}\right)$	$R\ \left(\frac{3}{4} - \frac{q}{4}\right)$
$T\ (1)$		
$B\ (0)$		

$\Delta k = 4 \,(\mathrm{mod}\, 4)$:

$x = 320$	$L\ \left(\frac{q_l}{2}\right)$	$R\ \left(1 - \frac{q_l}{2}\right)$
$T\ \left(\frac{3}{4} + \frac{q_l}{4}\right)$		
$B\ \left(\frac{1}{4} - \frac{q_l}{4}\right)$		

As above, we can discard cases $\Delta k = 1$ and $\Delta k = 2 \,(\mathrm{mod}\, 4)$ based on the observation that the distribution of actions in population 2 is not uniform. With $q_l \geq 1/8$, cases $\Delta k =$

$1 \pmod 4$, $3 \pmod 4$ entail at least 37.5% of population 2 playing L. The data only show a 16%, hence we discard this possibility as well. The only case that is left therefore is $\Delta k = 4 \pmod 4$, which yields the same distribution of actions for population 2 as case 2 above (that is, $q_l/2$.)

Overall, we are left with two possibilities, both entailing that a fraction $q_l/2$ of population 2 plays L. If we choose q_l to match the empirical distributions, we obtain $q_l = 0.32$, which falls precisely in the interval $(1/8, 1/2)$. The only explanation that appears consistent with the empirical distribution of population 2 therefore is the following, which corresponds to the case $\Delta k = 4 \pmod 4$:

Calibration: $q_l = 0.32$ [Data in Brackets]

$x = 320$	L $(16)^*$	R $(84)^*$
T (83) [96]		
B (17) [4]		

We next consider the case $x = 44$, maintaining that $q_l \in (1/8, 1/2)$ from the previous exercise. Notice that for this game, player 1 plays B as soon as he attaches probability at least $1/11$on R being played. First, it is easy to show that for any increasing cost functions c_l, c_h, there exists $x > 40$ sufficiently low that the both types of population 1 would choose B at their cognitive bound. Assuming that $x = 44$ is 'sufficiently low', a reasoning similar to the one presented in Section 5.2.2 delivers the following results: all low types of population 1 play B, while the low types of population 2 are uniformly split; if $q_l \in (1/8, 1/2)$, the 50% of high types of population 1 that believe that the low types of population 2 play R will play B (because $q_l > 1/8 > 1/11$), and the 50% of high in population 2 that anticipate this will play L. The remaining 50% of high types in population 1 play according to their own cognitive bound, that is B. Since the high types in population 2 have the same cost function as the latter ones, but higher incentives, they would be able to anticipate this, and respond playing L. Hence, all high types in population 2 play L.

Summarizing, for $q_l \in (1/8, 1/2)$ our findings for the three games are:

$x = 80$	L$\left(\frac{1}{2}\right)$	R$\left(\frac{1}{2}\right)$
T $\left(\frac{1}{2}\right)$		
B $\left(\frac{1}{2}\right)$		

$x = 320$	L $\left(\frac{q_l}{2}\right)$	R $\left(1 - \frac{q_l}{2}\right)$
T $\left(\frac{3}{4} + \frac{q_l}{4}\right)$		
B $\left(\frac{1}{4} - \frac{q_l}{4}\right)$		

$x = 44$	L$\left(1 - \frac{q_l}{2}\right)$	R$\left(\frac{q_l}{2}\right)$
T (0)		
B (1)		

D.2 Coordination Game with a Secure Option

This game has two pure-strategy Nash equilibria, (L, L) and (H, H), which are not affected by the value of x. Hence, anchors equal to (L, L) or (H, H) would generate a path of reasoning in which respectively L or H is repeated. Anchors (L, H) or (H, L) determine a cycle alternating between H and L, which is also independent on the value of x. The paths generated by anchors that involve S, instead, vary with the value of x, but since action S is dominated, it is never part of any path of reasoning for any $k > 0$. Nonetheless, it shapes player 1's incentives to reason, as an increase in x changes the value of doing a step of reasoning when player 1 is in a state in which action H is regarded as the most sophisticated. Applying equation (8) to this game, with payoffs parameterized by x, we obtain the following value of reasoning functions:

$$v_1(k) = \begin{cases} 180 & \text{if } a_1^{k-1} = L \\ \max\{90, x\} & \text{if } a_1^{k-1} = H \end{cases} \tag{11}$$

$$v_2(k) = \begin{cases} 90 & \text{if } a_2^{k-1} = H \\ 180 & \text{if } a_2^{k-1} = L \\ 140 & \text{if } a_2^{k-1} = S \end{cases} \tag{12}$$

Similar to the asymmetric matching pennies games discussed above, any path in which agents cycle between action L and action H induce a v_1 function that alternates between 90 and 180. Whether the spikes are associated to odd or even k's depends on the anchor. When $x = 400$, the incentives to reason do not change for player 2, but $v_1$changes alternating between 180 and 400: the 'spikes' at 400 replace what would be 'troughs' at 90 with $x = 0$.

The experimental results show that 96% of player 1 and 84% of player 2 played H when $x = 0$. One possible explanation is that in the baseline coordination game the efficient equilibrium is sufficiently focal that most individuals approach the game with $a^0 = (H, H)$ as an anchor.[35] Under the assumption that anchors are uniformly distributed, the only way that such a strong coordination on H can be explained is by assuming that the 'spikes' and 'troughs' determined alternating between 180 and 90 are already sufficiently pronounced that the low types involved in a reasoning process that determines a cycle stop their reasoning at H. Hence, with $x = 0$, agents that approach the game with anchors $a^0 = (L, L)$ play L, all others play H (because they either settle on a constant H, as in $a^0 = (H, S), (H, H)$, or they determine a cycle, as in $a^0 = (H, L), (L, H), (L, S)$). The predictions of the model therefore are the following:

$x = 0$	L (1/6)	H (5/6)	S (0)
L (1/6)			
H (5/6)			

We next consider the case $x = 400$, maintaining the assumption that the anchors are uniformly distributed. For the same reasons discussed above, for any pair of (increasing) cost functions c_l, c_h, there is x sufficiently high that all low types of population 1 with a reasoning process that involves a cycle stop at L. If $q_l < 2/3$, however, this is not enough to induce the high types of population 2 to play L as well. Hence, if we insist on $q_l = 0.32$ calibrated above, both the low and the high types in population 2 play according to their own cognitive bound. Since the incentives to reason were not affected by the change in x for these individuals, the assumptions above entail that they play H. Hence, in population 2, all individuals with anchors $a^0 \neq (L, L), (L, S)$ play H, the others play L. It remains to consider the high types of population 1. Since with $x = 400$ these types have stronger incentives to reason than the high types of population 2, any of these types involved in a cycle anticipates that

[35]The change in behavior observed when $x = 400$ could then only be explained by arguing that this payoff transformation changes the way the agents approach the game. While we think this is a plausible explanation, we maintain the assumptions that anchors are uniformly distributed, and explore to what extent the mere change in incentives may explain the observed variation in behavior, independent of the possible change in the anchors.

both types of population 2 would play H, hence they respond with H. Thus, in population 1, only the individuals whose anchor is $a^0 = (H, H)$ and the high types with anchors $a^0 \neq (L, L), (L, S)$ play H, that is a total of $1/6 + \frac{(1-q_l)}{2}$, or $2/3 - q_l/2$. The others play L. To determine the percentages of coordination in (L, L) and (H, H), we assume independence in the distributions of play between the row and the column players.
Summarizing:

$x = 400$	L (1/3)	H (2/3)	S (0)
L ($1/3 + q_l/2$)			
H ($2/3 - q_l/2$)			

with $q_l = .32$ calibrated from the matching pennies game:

$x = 400$	L (33)	H (67)	S (0)
L (49)	16		
H (51)		34	

Treatment	Own label	Opponent's label	Own payoffs	Opponent's payoffs	Own Tutorial	Opponent's Tutorial
Tutorial [D]	*I* (*II*)	*I* (*II*)	Low	Low	Yes	Yes
Mixed tutorial [E]	*I* (*II*)	*II* (*I*)	Low	Low	Yes	No
Mixed tutorial [F]	*I* (*II*)	*II* (*I*)	Low	Low	Yes	No

Table 10: Treatment summary for post-tutorial treatments [D], [E] and [F]

Supplementary Materials

E Additional Treatments

In this Appendix, we describe additional treatments conducted on a subset of the subjects with the aim of exploring less immediate predictions of our theory as well as possible directions for future research. These designs are more complex and cognitively demanding for the subjects, and some rely on the theoretical model more than they directly test it. Overall, the results are encouraging, especially in light of these factors. The instructions of these additional treatments are in Appendix A.1.3 (treatments [D], [E] and [F]) and Appendix A.1.4 (treatments [K] and [L]).

E.1 Identifying Beliefs

The theoretical model of Section 2 offers a clear distinction between a player's cognitive bound, $\hat{k}_i$, and his behavioral k_i, which is determined by his beliefs about the opponent, k_j^i. Recall that we have made the simplifying assumption that subjects view label I opponents to be more sophisticated than they are themselves. As previously mentioned, this condition is not necessary for our theoretical predictions. Here, we present a test of whether players overall do indeed play according to their bound $\hat{k}$ when playing label I opponents. As we demonstrate below, the evidence seems consistent with this assumption.

Specifically, we aim to test whether subjects play according to their own cognitive bound in treatments [A] and [B] or whether they are responding to their beliefs about the opponents' cognitive bound. In the context of our model, this requires a mechanism for setting the players' own costs to zero while holding their beliefs about their opponents constant. For instance, consider the example of Figure 2 (p. 14). If we change the cost function of player i so that $c_i(k) = 0$ for every k, then player i would play according to $k_j^i = \bar{k}_j^i$. If the cognitive bound had not been binding before (as in Figure 2.a), then the agent's behavior would remain the same. If instead his cognitive bound had been binding beforehand (Figure 2.b), then his behavior may change, since $k_j^i < \bar{k}_j^i$ before the decrease in cost. Hence, within our model, the choice of a lower number (higher k) suggests that the player's action had been determined by his cognitive bound rather than his beliefs. Our treatments [D], [E] and [F], summarized in Table 10, are designed precisely to operationalize this thought experiment.

After having administered the main treatments, we expose all eighty subjects from four of the six sessions (two for the endogenous and two for the exogenous classifications) to a

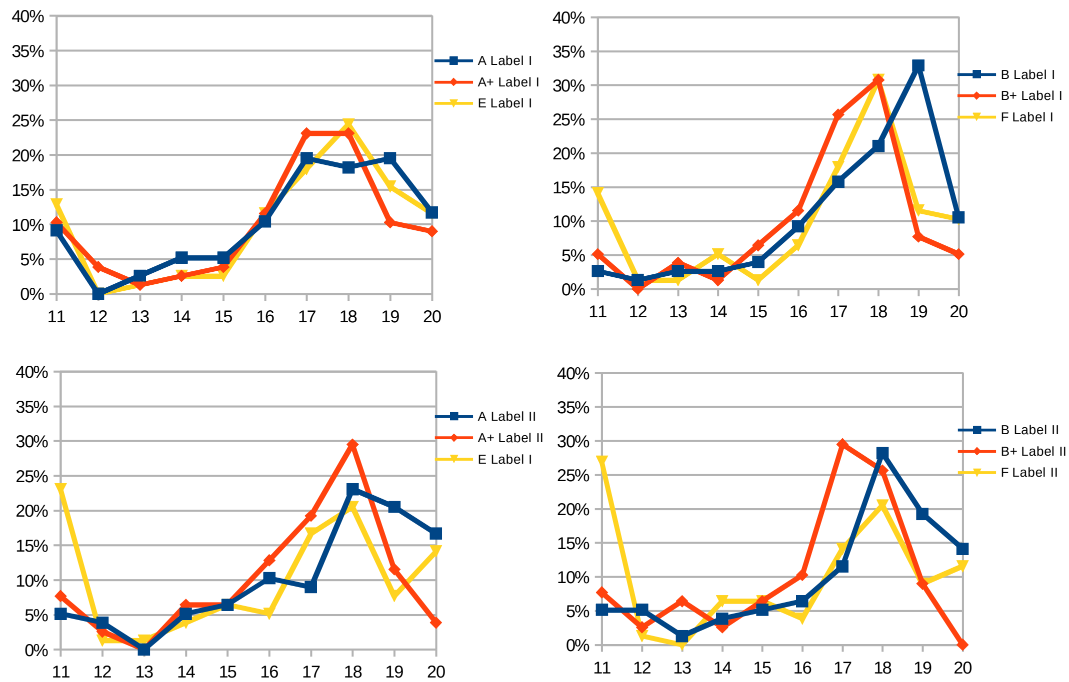


Figure 10: Post-tutorial treatment [E] compared to pre-tutorial treatments [A] and [A+]; post-tutorial treatment [F] compared to [B] and [B+] for labels I and II; frequency distributions

'game theory tutorial'. This tutorial explains how, through the chain of best replies, 'infinitely sophisticated and rational players' would play $(11, 11)$. We interpret offering this explanation as setting the subjects' cognitive costs to zero. We then proceed with three new (post-tutorial) treatments, each repeated twice. In treatment [D], we instruct each subject to play the baseline game (with low payoffs) against another subject who has also been given the same tutorial. Not surprisingly, a high fraction of the subjects (48% of label I and 55% of label II) announce 11. In treatment [E], we instruct the subjects who had previously received the tutorial to play the baseline game against a player of the same label who had *not* received the tutorial (that is, as in the homogenous treatment [A]). Analogously, treatment [F] contains the same instructions but with the subjects facing an opponent from a different label (as in [B]). Hence, subjects essentially face the same opponents in treatments [E] and [A] (and in [F] and [B]), but their cost function has been 'turned off' in treatments [E] and [F]. If their cognitive bound are *not* binding in (pre-tutorial) treatments [A] and [B], then the distributions of actions in (post-tutorial) treatments [E] and [F] should be the same as in [A] and [B], respectively. Figure 10 displays the results for the two labels. Comparing [A] to [E] and [B] to [F], we observe that the distributions of actions shift to the left, with different degrees across labels and treatments. Through the lens of our model, the data suggest that the cognitive bound for label I in treatment [A] and for label II in treatment [B] are binding for at least some of the agents.

We emphasize, however, that this interpretation should be taken with a grain of salt. It is not obvious that a highly manipulative intervention such as providing the tutorial would change players' cost function without affecting their beliefs. Nonetheless, given the complexity of the design and the instructions, we find these results to be encouraging.[36]

Figure 10 also shows the distributions for the treatments with high payoffs, [A+] and [B+]. This figure reveals how, as predicted by our model, increasing the incentives (from [A] to [A+] and from [B] to [B+]) produces effects analogous to reducing the cost function (from [A] to [E] and from [B] to [F]). Unlike treatments [E] and [F], however, subjects in treatments [A+] and [B+] play against opponents who also have higher incentives than in [A] and [B]. Thus, treatments [A+] and [B+] cannot be directly compared to [E] and [F].

E.2 Reasoning about opponents' incentives

In the design of treatments [A+], [B+] and [C+], relative to [A], [B], [C], we increase the payoff for undercutting the opponent for both players in the game. Thus, the shifts in the distributions towards lower numbers observed in Section 4.1 may conflate two distinct effects. The first effect is the possible increase in the cognitive bound of player i, and the second is the change in i's beliefs about j's cognitive bound due to the change in j's incentives. Both effects would determine an increase in the behavioral k_i, hence a shift of the distribution towards lower actions.

To illustrate these effects using our model, consider the example in Figure 2.b (p.14). When $v_i = v_j = v$, the cognitive bound and behavioral level of play are $\hat{k}_i = k_i = 3$. Now, suppose that v_i is increased up to v^*, while holding fixed $v_j = v$. In this case, $\hat{k}_i = 6$ and $\hat{k}^i_j = \bar{k}^i_j = k^i_j = 4$. Player i's response is to play according to $k_i = 5$, which is higher than the original level of 3. Suppose next that v_j is also increased to $v_j = v^*$. Then player i's cognitive bound becomes binding, and k_i increases to $k_i = \hat{k}_i = 6$. The movement from 3 to 5 is thus due to the increase in $\hat{k}_i$ alone, induced by an increase in v_i; the further change from 5 to 6 instead is determined by i's reasoning about the change in his opponent's incentives.[37]

The following treatments, summarized in Table 11, are aimed at testing whether subjects in our experiment reason about their opponents' incentives independently of their own. In a sense, the exercise is of a similar spirit to treatment [C], in which subjects play against the number chosen by an opponent in treatment [B]. Similarly, in treatments [K] and [L] agents play the high-payoff game against the number chosen by an opponent in treatments [A] and [C], respectively. Both treatments are administered after the main treatments to all forty subjects from two sessions (one exogenous and one endogenous), and each is repeated three times.

[36] Interestingly, a relatively high percentage of subjects from label II play 11 in post-tutorial treatments [E] and [F], compared to [A] and [A+]. Furthermore, the percentage of label II's who play 11 in [E] and [F] is nearly twice the percentage of label I's.

[37] In general, our model implies that $\hat{k}_i$ (weakly) increases whenever v_i is increased. Furthermore, if v_j is held constant, $\hat{k}^i_j$ increases only if $\bar{k}^i_j$ (the intersection between c^i_j and v_j) had been larger than $\hat{k}_i - 1$ in the first place. In Figure 2.a, for instance, increasing v_i without changing v_j does not affect $\hat{k}^i_j$, hence k_i. The opposite is true in Figure 2.b, where $\hat{k}^i_j$ increases as v_i is increased, until $\hat{k}^i_j = \bar{k}^i_j$ (that is, when v_i is sufficiently high that $\hat{k}_i \geq 5$.)

Treatment	Own label	Opponent's label	Own payoffs	Opponent's payoffs	Replacement of opponent's opponent
Mixed payoffs-homogeneous [K]	I (II)	I (II)	High	Low	No
Mixed payoffs-heterogeneous [L]	I (II)	II (I)	High	Low	Yes

Table 11: Treatment summary for mixed-payoffs treatments [K] and [L]

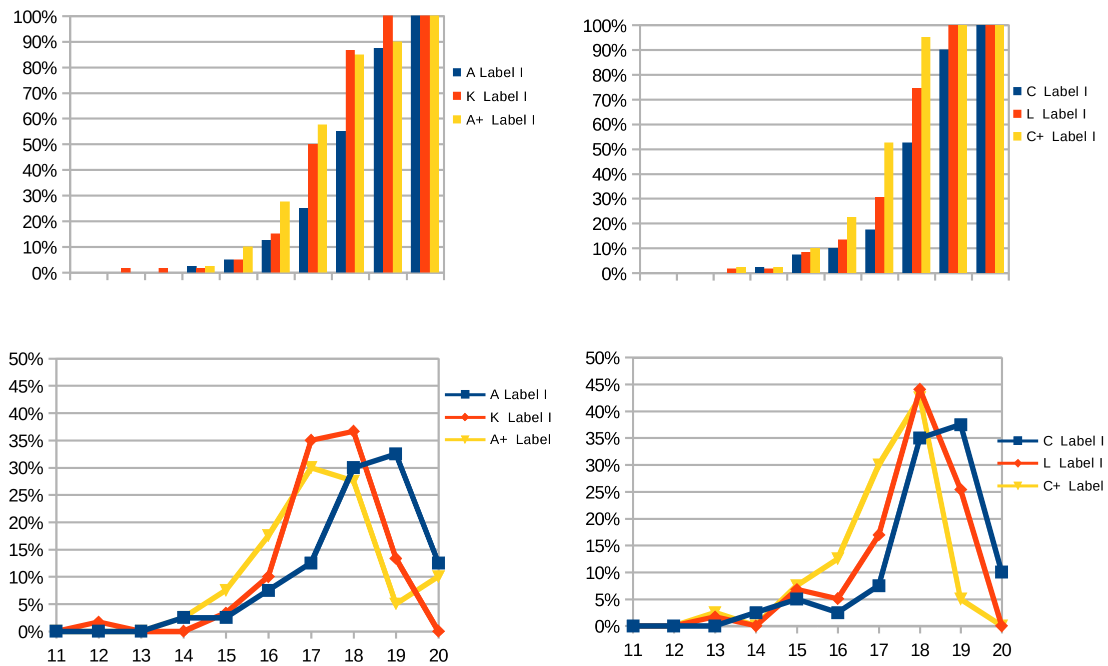


Figure 11: Treatment [K] compared to [A] and [A+]; treatment [L] compared to [C] and [C+] for label I; cumulative distributions (top) and frequency distributions (bottom)

These treatments add a further layer of complexity, since the individual is told in treatment [K] (resp., [L]) that he is playing the high-payoff game against the number chosen by an opponent of the same (other) label himself playing the low payoff game against opponent of the same (other) label. Treatment [L] is especially complex: for player i, both the payoffs and the label of i's opponent *and* of the opponent's opponent are different from i's own payoff and label.

By comparing treatments [K] and [L] with treatments [A] and [C] and with treatments [A+] and [C+], we can, in principle, disentangle the two effects mentioned above. The shift from [A] to [K] (and from [C] to [L]), due solely to the increase of each subject's own payoffs and not his opponent's, may be attributed to the increase of subjects' own cognitive bound. It should be observed only if the cognitive bound in treatments [A] and [C] had been binding (see footnote 37); the further shift from [K] to [A+] (and from [L] to [C+]) instead can be imputed to the increase in subjects' beliefs about their opponents' behavior due to the increase of their payoffs.

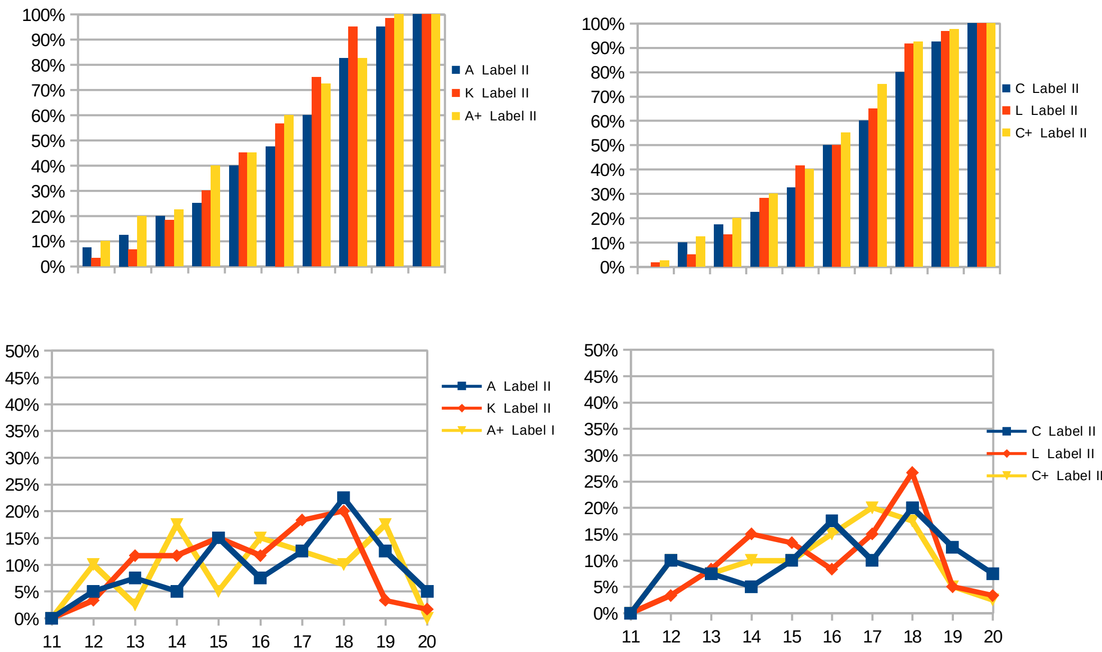


Figure 12: Treatment [K] compared to [A] and [A+]; treatment [L] compared to [C] and [C+] for label II; cumulative distributions (top) and frequency distributions (bottom)

Figures 11 and 12 show the results of these treatments for labels I and II, respectively. The results are roughly in line with the predictions of the theory. The empirical distribution of [A] first order stochastically dominates [K] everywhere for label I other than at 14, and in most of the curve for label II. Distribution [K] first order stochastically dominates [A+] in the majority of the curve, although this appears more tentative. The results for [L], however, are surprisingly clean: the distribution of [L] lies ‘in between’ the distributions of [C] and [C+] nearly everywhere for label I and in most regions for label II. Furthermore, for label I, the theory predicts that the increase in k_i from [C] to [L] should be at most one, which seems roughly confirmed by the small shift in distribution from [C] to [L]. In this case, and consistently with the theory, the movement from [C] to [C+] for label I is mainly due to the increase in the opponents’ payoffs, and not solely to the agent’s own incentives. In light of the complexity of these treatments and the difficulty of the instructions, these results appear to be surprisingly good. But these factors suggest caution in interpreting the results.