Magic Mirror on the Wall, Who Is the Smartest One of All?*

Yoram Halevy† Johannes C. Hoelzemann‡ Terri Kneeland§

September 3, 2025

Abstract

In the leading model of bounded rationality in games, each player best-responds to their belief that the other players reason to some finite level. This paper investigates a novel behavior that could reveal whether the player's belief lies outside the iterative reasoning model. This encompasses a situation where a player believes that their opponent can reason to a higher level than they do. We propose an identification strategy for such behavior, and evaluate it experimentally.

JEL Classification: C72, C92, D91.

Keywords: Bounded rationality, higher-order rationality, level-k, cognitive-hierarchy, game theory, equilibrium, rationalizability, preference elicitation.

*We thank Jose Apesteguia, Ayala Arad, Gabriel Carrol, Vincent Crawford, Rahul Deb, Amanda Friedenberg, Li Hao, Philippe Jehiel, Rosemarie Nagel, Antonio Penta, Marcin Peski, Mike Peters, Doron Ravid, Anne-Katrin Roesler, Ariel Rubinstein, Marciano Siniscalchi, Colin Stewart, and Georg Weizsäcker for very helpful discussions and are grateful to participants in the following conferences, seminars and workshops for comments and suggestions: ANZWEE 2025 (Sydney), Barcelona School of Economics Summer Forum 2022, Bar-Ilan University, Bounded Rationality workshop 2022 (Tel Aviv), CUHK-HKU-HKUST Theory seminar, D-TEA 2022 (Paris), ESA North American winter meetings, Florida State University, Hebrew University of Jerusalem, Koç Üniversitesi, M-BEES 2025, Nazarbayev University, Simon Fraser University, UC-SFU-UBC Theory seminar, UEA Behavioral Game Theory Workshop 2023 (Norwich), Universidad de Alicante, Universität Duisburg-Essen, Universität Wien, Université de Montréal, University College London, University of California at Davis, University of Chicago, University of Ottawa, University of Texas at Dallas, University of Toronto, SAET, SITE (Psychology & Economics) 2022, and Wirtschaftsuniversität Wien. The experiment was conducted via the Toronto Experimental Economics Laboratory in April 2020, under The University of Toronto Research Ethics Board 083 approval #00037554. Financial Support from SSHRC as well as ERC Grant SUExp - 801635 are gratefully acknowledged.

†University of Toronto & Hebrew University of Jerusalem: https://uoft.me/yoramhalevy & yoram.halevy@utoronto.ca.

‡Universität Wien: http://www.johanneshoelzemann.com & jc.hoelzemann@gmail.com.

§University College London & University of Calgary: http://www.tkneeland.com & t.kneeland@ucl.ac.uk.

1 Introduction

The leading models of bounded rationality in games, as level-k and cognitive hierarchy, are iterative 'top-down' models of reasoning: a player with a finite level of reasoning believes others can reason to a strictly lower level and best responds to that belief. This restriction is critical in how the model is operationalized – it ensures that a player requires only a finite number of steps of reasoning to optimally respond to their belief. Importantly, a player who can do k steps of iterated reasoning (i.e., k steps of "I think, you think, I think, ...") can only model others as being capable of doing at most $k-1$ steps of iterated reasoning.[1] This ability to model the behavior of others is a key assumption in these models. This, however, leads to a natural and interesting question: what happens if a player believes others may reason to a higher level than they do? For example, how will a player respond if they believe that their opponent is more sophisticated than them?

We propose a behavior that reveals to an analyst that Ann, who is playing a game with Bob, is reasoning about Bob's behavior outside of the iterative 'top-down' model of reasoning. We then implement a novel experimental design that allows us to identify this behavior experimentally and evaluate its pervasiveness in the population. We also investigate whether Ann's behavior depends on Bob's observed characteristics that may be correlated with his sophistication.

Recall that in iterative 'top-down' reasoning models players' beliefs are anchored in the behavior of a specific non-rational L0 type, and types are heterogeneous in their level of reasoning. The L1 type performs one level of reasoning and best responds to the L0 type. In turn, the L2 type performs two levels of reasoning and best responds to some belief over L0 and L1 types, and so on with the Lk type best responding to some belief over L0, ..., L$(k-1)$ types. But how would Ann behave if she believed that Bob may be more sophisticated than her? Within the prism of the iterative 'top-down' model of reasoning, it implies that

[1] Any player who can reason about their opponent doing m steps must necessarily be able to do at least $m+1$ steps of reasoning themselves.

although she would believe that Bob is rational (since she is rational), she will not be able to model his behavior. Still, Ann's behavior would be consistent with 2-rationalizability, which allows all actions that are consistent with rationality and belief in others' rationality.

We design two diagnostic games that allow the analyst to identify this behavior. The first is a *dominance-solvable* game ("DS") in which Bob has a dominant strategy. This game permits the analyst to identify if Ann "believes that Bob is rational." Using the second game – which we refer to as the *iterative-reasoning* game ("IR") – the iterative 'top-down' model of reasoning *together* with belief in rationality makes the sharp prediction that Ann would value IR strictly more than DS. However, if Ann only believes that Bob is rational, but her reasoning process is not captured by the model (but is consistent with 2-rationalizability), she may value DS more than IR. Importantly, these inferences do not depend on Ann's risk or social preferences. This results in a conservative estimate of the proportion of participants who are inconsistent with the iterative 'top-down' model of reasoning.

Our identification strategy uses a more general anchor than the standard L0 type. We consider a rational, but non-strategic, L1 type to anchor the iterative 'top-down' model of reasoning. This player concentrates only on their own payoff, without making any strategic considerations. This increases the set of possible actions that are consistent with the L1 type, includes the "standard" L1 type (who best-responds to uniform play of the L0 type), and accommodates other focal behaviors.

Our test to identify if Ann's behavior is consistent with the prediction of a generalized iterative reasoning model may be extended to the case where Ann may not believe that Bob is rational, if the form of irrationality considered is a random choice of action by Bob (a uniform play by the L0 type, as is typical in many models). In this case, the ranking of the DS and IR games is unaltered.

The novel experimental design that we employ has four components. The first are the two diagnostic games: IR and DS. The second are two control games that rule out other confounding factors that can contribute to preferring DS over IR. Third, we investigate whether participants' reasoning process (iterative 'top-down' models of reasoning or 2-rationalizability) depends on their opponents' observed characteristics. To achieve this, we exogenously vary the participants' opponent type: they face either a Ph.D. student in Economics or an undergraduate student of any discipline. The fourth component is a preference-elicitation mechanism over the games. Rather than directly eliciting a choice between the two diagnostic games, participants first choose their actions in each game (and against each potential opponent), and then we elicit their respective *valuations*.[2,3] This allows the analyst to infer both participants' preferences between the two diagnostic games and participants' (confidence in their) beliefs about their opponents' behavior. Moreover, we can exploit the valuation data to isolate those participants who believe that their opponent is rational, as the predictions in our games are the starkest for this subset of participants.

We find that approximately half of the choices made by participants are inconsistent with the iterative 'top-down' model of reasoning, especially for those who believe that their opponents are rational – where the model's predictions are inconsistent with 64% of choices. Moreover, approximately 72% of participants exhibit a stable model of reasoning irrespective of the opponent's characteristics. Among the remainder, the results are split: roughly 12% make choices consistent with iterative 'top-down' reasoning against an undergraduate but not against a Ph.D. student, while roughly 16% exhibit the opposite pattern.

Pioneering scholarly contributions in the iterative 'top-down' reasoning literature include Nagel (1995), Stahl and Wilson (1994; 1995), Costa-Gomes, Crawford, and Broseta (2001), Camerer, Ho, and Chong (2004), and Costa-Gomes and Crawford (2006). For a review of this literature, see Crawford, Costa-Gomes, and

[2]Heinemann, Nagel, and Ockenfels (2009), Coricelli and Nagel (2009), and Nagel, Brovelli, Heinemann, and Coricelli (2018) use a related strategy to elicit certainty equivalents in coordination games; however, in their context, the elicited valuations affect both the payoffs in the games and their value.

[3]To allow participants to recall their reasoning in the valuation stage, we encouraged them to write it down in a text box. We use this information to gather further qualitative evidence on their choice process.

Iriberri (2013). By construction, these papers do not consider the questions we investigate here.

Arad and Rubinstein (2012a) and Kneeland (2015) developed novel experimental designs to identify levels of reasoning in an iterative model. Moreover, in the former design, the authors explicitly asked participants about their thought process when making their choices to gain a better understanding of participants' behavior. Arad (2012) proposed a new allocation game to study iterative reasoning and the performance of the level-k model, and showed that level-k thinking accounts for a smaller number of choices made by participants than in other experiments. Further, Arad and Rubinstein (2012b) studied how participants reason iteratively on few dimensions, or features, in an allocation game (Colonel Blotto). Subsequently, Arad and Penczynski (2024) studied some other environments of resource allocation with communication between participants, and confirmed that many participants engage in multi-dimensional iterative reasoning.

Also related to our work is Agranov, Potamites, Schotter, and Tergiman (2012) who manipulated participants' beliefs about the cognitive levels of the players they are playing against; and Alaoui and Penta (2016) who studied a model of iterative reasoning where player's depth of reasoning is endogenously determined. More recently, Alaoui, Janezic, and Penta (2020) further developed an experimental design strategy to distinguish level-k behavior driven by participants' beliefs from their cognitive bounds, and found an interaction between participants' own cognitive bound and reasoning about the opponent's reasoning process. Gill and Prowse (2016) investigated how cognitive ability and character skills influence the evolution of play in repeated strategic interactions and estimate a structural model of learning based on level-k reasoning. Georganas, Healy, and Weber (2015) examine whether the level-k model generates reliable cross-game testable predictions at the individual player level, and find that observed levels are mostly consistent within one family of similar games, but not across families of games.

Our work also draws on the epistemic game theory literature, particularly in how it approaches strategic uncertainty – that is, uncertainty about the play of others. A key element of our experimental design is the use of the 2-rationalizable solution concept as our alternative model of behavior, i.e., how people behave when they are reasoning outside of the iterative 'top down' model.

In the iterative top-down approach, strategic uncertainty arises indirectly in three ways. First, through assumptions about L0 behavior (e.g., if L0 types are assumed to choose actions uniformly at random, then L1 types face uncertainty about others' play). Second, through uncertainty about others' reasoning levels; and third, through uncertainty about others' risk preferences. For the latter, if players differ in their levels of reasoning or in their risk preferences, their optimal actions may differ, implicitly generating strategic uncertainty about the play of others.

In contrast, the epistemic approach models strategic uncertainty explicitly, restricting it only through epistemic conditions such as rationality and beliefs in others' rationality. The 2-rationalizable solution concept captures this by allowing a wide range of beliefs consistent with these conditions, rather than relying on *ad hoc* assumptions about L0 behavior, levels of reasoning, or preference heterogeneity.

Our experimental design is constructed to isolate and control for the three sources of strategic uncertainty accounted for by the iterative 'top down' model of reasoning. This allows us to identify behavior that falls outside the predictions of the iterative 'top down' model but is consistent with 2-rationalizability. Put differently, our design allows us to observe behavior that reflects the broader and more explicit conception of strategic uncertainty embedded in the 2-rationalizable model, but not captured by the iterative 'top-down' model of reasoning.

Our findings highlight the importance of the epistemic approach. They suggest that explicitly modeling strategic uncertainty may capture observed behavior that implicit treatments via modified iterative 'top-down' models are inconsistent with. This complements earlier works that contrast the explicit approach of epistemic game theory and other implicit modeling approaches across domains, including auctions (Battigalli and Siniscalchi 2002, Dekel and Wolinsky 2003, Kosenkova

2019), bargaining (Friedenberg 2019), and identifying levels of reasoning (Brandenburger, Friedenberg, and Kneeland 2020).

The paper proceeds as follows. Section 2 introduces the design and the set of diagnostic games as well as the two control games. It builds the theoretical background necessary for our experiment – discussed in Section 3 – and the identification strategy used in the analysis carried out in Section 4. Section 5 offers a more formal analysis. Finally, Section 6 concludes with a brief discussion of the results. The Appendix contains further analyses, details on participants' individual behavior, the experimental instructions, and screenshots of the experimental interface.

2 The Design

We employ both an iterative 'top-down' model of reasoning, based on level-k and cognitive hierarchy, and the solution concept of 2-rationalizability to guide our experimental design, identification strategy, and analysis. We provide a brief description of the model and the concept here and engage in a discussion on how these interact with our setup in the next subsection. A more formal and general analysis is provided in Section 5.

2.1 Building Intuition: Model and Solution Concept

Iterative 'top-down' model of reasoning In this model, players anchor their beliefs in a naïve model of others' behavior and adjust their beliefs by a finite number of iterated best-responses. To date, these models have been anchored in an "irrational" (L0) player-type who either plays each strategy with equal chance or chooses some salient action, depending on the application. Players of level-k ($k > 0$) are rational in the sense of best-responding to their beliefs, but players of different k differ in their beliefs over the action(s) played by their opponents.

We consider a more general model of reasoning, with a different cognitive interpretation of L1. Instead of conducting a "standard" level-k analysis, our goal is to give the iterative 'top-down' approach the best possible chance of success by considering the most general model. Our model is anchored in the behavior of a non-strategic L1 type who makes decisions based solely on their own-payoff information. To build intuition for this type, consider a decision maker who chooses an action to allow for the possibility of achieving the highest possible payoff in a given game, or alternatively, chooses an action to maximize their average payoff. In both cases, the decision maker is non-strategic as they *never* form beliefs about their opponents' behavior. Nevertheless, their behavior may very well reflect their *own* payoff information and primary focus therein. If one views their choice of action independently of the strategic environment, L1-choices could be viewed as "rational." Since there are many possible criteria a decision maker could employ to determine their action choice, selecting an action in order to ensure the maximum or the average payoff being just two examples, we will use a partial-order approach to formalize this behavior.[4] Effectively, as long as an action is optimal under some own-payoff criteria, we would allow our non-strategic type (L1) to play it.[5]

Since we want to capture all reasonable own-payoff criteria that our decision maker could use, the only assumptions we impose are that the criteria must be non-strategic in nature, and respect the notion that higher payoffs are preferred, i.e., strict monotonicity. Consider two payoff vectors $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, \ldots, y_n)$ such that $\boldsymbol{x}$ is greater than $\boldsymbol{y}$; that is, $x_i \geq y_i$ for all $i \in \{1, \ldots, n\}$ with strict inequality for at least one i. In this case, it seems clear that $\boldsymbol{x}$ should be preferred to $\boldsymbol{y}$ if our decision maker prefers higher payoffs. Further, since we are trying to capture the behavior of a non-strategic type, we should ignore any information contained in the ordering of the payoff vectors, as any concerns for ordering would reflect strategic considerations. Thus, we propose the following partial order $\succ_1$: $\boldsymbol{x}$ is preferred to $\boldsymbol{y}$ if there exists a permutation of $\boldsymbol{x}$ that is

[4]Text data collected indicates that "average" and "maximum" payoff are terms with relatively high Term Frequency-Inverse Document Frequency scores, a numerical statistic that is intended to reflect how important a word is to a comment in a collection.

[5]Coricelli and Nagel (2009) as well as Nagel, Brovelli, Heinemann, and Coricelli (2018) found that players who do not engage in high-level strategic thinking have similar brain activation to decision makers who make risky decisions in non-strategic environments, providing physical support to our typology of L1 as rational but non-strategic.

greater than $\boldsymbol{y}$. We then allow our non-strategic type to play any action that is undominated according to $\succ_1$.

Notice that the binary relation $\succ_1$ is not, in general, complete. For example, consider two payoff vectors $\boldsymbol{a} = (20, 0, 10)$ and $\boldsymbol{b} = (12, 8, 16)$. Here, neither $\boldsymbol{a}$ is preferred to $\boldsymbol{b}$ nor $\boldsymbol{b}$ is preferred to $\boldsymbol{a}$. This reflects the fact that strategy a might be optimal under one criteria (e.g., it has the highest payoff), yet strategy b might be optimal under another criteria (e.g., it has the highest arithmetic mean).[6] Alternatively, consider the two payoff vectors $\boldsymbol{c} = (20, 9, 14)$ and $\boldsymbol{d} = (12, 8, 16)$ that are comparable according to $\succ_1$; that is, $\boldsymbol{c}$ is preferred to $\boldsymbol{d}$.

In general, the partial order $\succ_1$ incorporates many potential own-payoff heuristics that seem both intuitive and reasonable. The set of actions an L1 type will choose from – the actions that are undominated through $\succ_1$ – must always contain an action that leads to the highest payoff, an action with the highest minimum payoff, as well as the action with the highest arithmetic mean.[7] Further, notice that the action with the highest arithmetic mean is equivalent to the action that maximizes a player's expected payoffs under the belief that others' play each action with equal probability. As such, our approach nests the standard level-k and cognitive hierarchy models as a special case as they typically assume that the L0 type plays uniformly random.[8]

The behavior of all higher types is then anchored in the behavior of the L1 type. A level-2 (L2) type assumes that all other players are the L1 type and chooses accordingly a strategy that maximizes their expected utility under some probability distribution over L1 strategies.[9] A level-3 (L3) type assumes that all other players are either L1 or L2 types and chooses a strategy that maximizes their expected utility under some probability distribution over both L1 and L2

[6] Note that probabilistic beliefs on the actions chosen by others, as is assumed in the literature to date, induces a complete ranking on the player's actions.

[7] All three of these own-payoff heuristics were shown to have explanatory value as part of a focal L0 type in Wright and Leyton-Brown (2014).

[8] Moreover, our approach also nests many special cases of non-strategic behavior proposed in the level-k literature to express notions of 'focal points' such as playing 20 in Arad and Rubinstein (2012a)'s 11-20 game. Hence, in the current setup, the L1 type will play that strategy but beyond relabelling of levels – nothing will change.

[9] Most iterative reasoning applications assume that players are risk-neutral and hence maximize expected payoffs. Importantly, we allow instead for *any* expected-utility preferences.

strategies. This process continues for higher-level types *ad infinitum* and, more generally, with Lk types choosing a strategy that maximizes expected utility given some belief over the play of strictly lower types.

2-rationalizability This solution concept can be intuitively understood via its relationship with the notion of rationality and reasoning about rationality. A player is *rational* if they play a best-response – maximize expected utility – given their subjective belief about how the game is played. A player *believes in rationality* if they believe that others are rational. That is, if they believe others are playing a best-response given their subjective beliefs about how the game is played. The solution concept of 2-rationalizable strategies incorporates both the assumption of rationality and belief in rationality.[10] The 2-rationalizable set is found by first finding the set of 1-rationalizable actions for each player. These are the actions played by a rational player: any action that maximizes a player's expected utility given some utility function and some belief about the play of others. The 2-rationalizable set comprises of all actions played by a rational player who believes others play actions in the 1-rationalizable set: any action that maximizes a player's expected utility given some utility function and some belief over the 1-rationalizable play of others. This solution concept is formally defined in Section 5.

Iterative 'top-down' model of reasoning and 2-rationalizability In the following, we highlight the relationship between the model and the solution concept introduced above. To start, notice that the iterative 'top-down' model of reasoning implicitly imposes assumptions about how types reason about rationality. We highlight three facts. First, all types with $k \geq 2$ are rational as they best respond to their beliefs about the play of others. Second, even though the L1 type cannot be considered rational in the game-theoretic sense as they are non-strategic and do not form beliefs about others' strategies, they nevertheless do play actions that

[10] The relationship between reasoning about rationality and k-rationalizable strategies follows from standard results, e.g., Bernheim (1984), Brandenburger and Dekel (1987), and Tan and da Costa Werlang (1988) among others.

are consistent with rationality. That is, any action that is undominated by $\succ_1$ is also a best response to some belief about others' play under some expected utility preferences. Third, the behavior of any Lk type with $k \geq 2$ is consistent with the assumption of belief in rationality. This result follows naturally since any such type believes that the behavior of others is, in fact, consistent with rationality.[11]

Further notice that the iterative 'top-down' model of reasoning imposes an additional assumption *beyond* reasoning about rationality. It imposes the assumption that beliefs are anchored in non-strategic play. Put differently, the L2 type cannot hold arbitrary beliefs about the play of the game. Rather, they must hold beliefs consistent with L1 play. While we use a generous definition of L1 play here to allow for a broad notion of non-strategic behavior, in many games this set of actions may still be small, even a singleton set. As such, one can interpret the L2 type here as a type that can model the play of others. Naturally, the same holds for higher levels. The L3 type that believes others are either L1 or L2 types cannot hold arbitrary beliefs about others' rational play, but rather must hold beliefs that are consistent with L1 or L2 play, and so on. Therefore, one can interpret the iterative 'top-down' model of reasoning as assuming that players in fact *can* model the play of others.

This is in sharp contrast to the concept of 2-rationalizability. This approach is grounded in the assumption that players can hold *any* beliefs about the play of others, and only requires those beliefs to be consistent with the assumption that others are rational. The assumption of rationality is less stringent than that imposed by L1 play. In this sense, 2-rationalizability can be interpreted as relaxing the assumption that players possess the ability to model the play of others, in contrast to iterative 'top-down' models of reasoning.

Key design assumptions In what follows, we will assume that players are strategic. For the iterative 'top-down' model of reasoning, this means that we will focus on the behavior of Lk types for $k \geq 2$ and not the non-strategic L1 players. This

[11]Notice that the model can easily be generalized if one wishes to allow for uncertainty over others' rationality by simply introducing an additional non-strategic type that randomizes uniformly over the set of actions. We shall discuss this in more detail in Section 5.

restriction is motivated by our main research question – whether players can model the play of others. This question is not applicable to non-strategic players who, by definition, do not reason about the play of others. Moreover, players who are rational and believe in rationality will play a key role in our design. As we assume that players themselves are rational since our focus is on types with $k \geq 2$, and investigate if they believe that others may be more sophisticated than them, it is natural to at least require them to believe that others are rational – even if they cannot model their behavior. As such, our design will make stark predictions for those participants who are rational and believe in rationality of others.

Recent work (Alaoui and Penta 2016; 2022, Alaoui, Janezic, and Penta 2022; 2025) demonstrates that individuals' levels of strategic reasoning can vary depending on the context or the structure of the game. This is consistent with the findings of Georganas, Healy, and Weber (2015), who also emphasize the sensitivity of reasoning levels to contextual factors. An important question then is: How sensitive is our design to the assumption that players' reasoning levels remain constant across the games? The answer is that our design is largely insensitive to such type changes across games. The only notable exception occurs when a player's type changes from an Lk type with $k \geq 2$ to a type with $k \leq 1$. Importantly, however, none of the above studies suggest that rational players become L0 under different game structures, so this would be an extreme version of having a wrong belief. Further, this kind of shift is not predicted by the endogenous depth of reasoning model applied in Alaoui and Penta (2016; 2022) and Alaoui, Janezic, and Penta (2022, 2025).[12] Thus, overall, we view our design as not dependent on the assumption that players' levels of reasoning remain fixed across games.

2.2 The Games

In order to identify behavior that reflects the player's belief that other players may be rational, but their behavior cannot be modeled, we judiciously designed two diagnostic games. One, where the ability to model the opponents' behavior

[12]For more details, see the discussion on page 39.

is important for how the participant values the game, and the other, where such an ability is less important.

The strategic form of these games is depicted in Figure 1.

IR

Player 1 \ Player 2	A	B	C	D
a	0, 12	12, 14	13, 12	11, 8
b	4, 0	14, 0	0, 16	6, 4
c	10, 16	0, 5	11, 0	12, 0
d	13, 7	8, 11	6, 10	0, 12

DS

Player 1 \ Player 2	A	B	C
a	0, 6	12, 3	11, 4
b	5, 15	13, 8	0, 13
c	12, 10	8, 9	0, 9

Figure 1: The Iterative-Reasoning Game (IR) and the Dominance-Solvable Game (DS).
In every cell, Player 1's payoff is displayed in the lower left, and the payoff to Player 2 is on the upper right.

The iterative-reasoning game "IR" The iterative 'top-down' model of reasoning predicts that Player 1 chooses an action in $\{a, b\}$ and Player 2 chooses an action in $\{B, C\}$. To gain intuition, consider first the simple case where for all $k \geq 2$, Lk type believes that the others are L$(k-1)$. Recall that Player 1 of type L1 considers their own payoffs but is non-strategic. This player chooses between the payoff vectors $\boldsymbol{a} = (0, 12, 13, 11)$, $\boldsymbol{b} = (4, 14, 0, 6)$, $\boldsymbol{c} = (10, 0, 11, 12)$, and $\boldsymbol{d} = (13, 8, 6, 0)$. Thus, the L1 type plays actions a or b, as actions c and d induce payoffs that are dominated by a permutation of a's payoffs. Either action a or b could be a natural focal action: action a is associated with the highest arithmetic mean, while action b is associated with the highest payoff. Similarly, Player 2 of level-1 plays action C. This action dominates all other actions according to $\succ_1$: it contains the highest arithmetic mean and highest payoff, and is therefore a natural focal action.

Any new iteration ("the next level") is a best response to the opponent's behavior. For example, the L2 type of Player 1 plays a and the L2 type of Player 2 plays B or C. Then, the L3 type of Player 1 plays a or b and the L3 type of Player 2 plays B. This process continues *ad infinitum*. Player 1's best responses are always in $\{a, b\}$, and Player 2's best responses are always in $\{B, C\}$.

The iterative 'top-down' model of reasoning is a more general model than this simple model. It explicitly allows players to hold arbitrary risk preferences within expected utility. Moreover, players may hold *any* belief about the expected-utility preferences of other players as well as over lower types L1, ..., L$(k-1)$ of other players. Even with these generalizations, it is still true that players will play actions in $\{a, b\}$ and in $\{B, C\}$. For details, see Section 5. As all strategic types (Lk where $k \geq 2$, i.e., those types that are rational and believe in rationality) of Player 1 in the generalized iterative 'top-down' model of reasoning play an action in $\{a, b\}$ and expect Player 2 to choose an action in $\{B, C\}$, their expected payoff must be *strictly greater than 12.*[13]

The solution concept of 2-rationalizability does not restrict Player 1 to value the game IR above 12. First, note that all actions of Player 2 in IR are 1-rationalizable, since for any of their actions there exists some belief about Player 1's play such that the action is a best response.[14] Second, if Player 1 believes that Player 2 is rational, they must believe that Player 2 plays a 1-rationalizable action. Such a player may reasonably hold *any* belief over the distribution of $\{A, B, C, D\}$. For example, Player 1 who believes that Player 2 is rational and assigns equal probability to all actions of Player 2 will choose the action a, and their expected payoff will be less than 12.

The dominance-solvable game "DS" The second diagnostic game is dominance-solvable in a single iteration, as A is a strictly dominant strategy for Player 2. It obviously dominates B and C according to $\succ_1$, as strict domination does not require strategic reasoning. That is, the L1 type and any higher type of Player 2 will play action A, which is a natural focal point for Player 2.

Now consider Player 1's behavior. If they are of level-1, they choose between payoff vectors $\boldsymbol{a} = (0, 12, 11)$, $\boldsymbol{b} = (5, 13, 0)$, and $\boldsymbol{c} = (12, 8, 0)$. Notice that a

[13] Player 1 may value IR exactly at 12. However, this can only occur with an extreme form of ambiguity aversion coupled with the player's set of priors including all degenerate priors. We elaborate on this point in Section 5 and document that it is not an empirical concern.

[14] Beyond B and C discussed above, A is a best-response to Player 1 playing c and D is a best response to Player 1 playing d.

permutation of $\boldsymbol{a}$ dominates $\boldsymbol{c}$, thus $\boldsymbol{a} \succ_1 \boldsymbol{c}$. However, neither $\boldsymbol{a} \succ_1 \boldsymbol{b}$ nor $\boldsymbol{b} \succ_1 \boldsymbol{a}$ is true. Either action a or b could be natural focal points for Player 1 of type L1. Action a is associated with the highest arithmetic mean, while action b is associated with the highest payoff. Since Player 2 of type Lk ($k \geq 1$) plays A, it must be that any Player 1 of type Lk ($k \geq 2$), best responds by playing c. From the above argument, it follows that the expected payoff of a rational Player 1 who believes that Player 2 is rational (all types with $k \geq 2$) equals 12.

In contrast to the IR game, the solution concept of 2-rationalizability *does restrict* the valuation of the DS game. Any player who is rational and believes in rationality must still behave exactly the same as in the iterative 'top-down' model of reasoning. Thus, any such player chooses action c and has an expected payoff of *exactly 12* irrespective of being an iterative-reasoner or not.

Player 1's preferences over IR and DS All players who are rational and believe that their opponents are rational prefer playing IR over DS in the iterative 'top-down' model of reasoning. The expected payoff of 12 in DS is strictly lower than the expected payoff in IR. As a consequence, a 'top-down' iterative-reasoner should strictly prefer to play IR over DS. However, a player who is rational and believes in rationality, yet falls outside the iterative 'top-down' model of reasoning, may very well prefer to play DS over IR. This behavioral difference is the core of our identification strategy.

Up to this point, we have restricted beliefs of rationality somewhat tightly for our strategic types (types with $k \geq 2$). In our iterative 'top-down' model of reasoning, there is no way for such a type to be uncertain about rationality; that is, there is no sense in which a type could believe others are playing actions that are not consistent with rationality. However, we can easily account for that by introducing a second non-strategic type that plays randomly, which we refer to as "level-0" ("L0 type"). We now simply permit a strategic Lk type to hold *any* beliefs over lower types $\{L0, L1, \ldots, L(k-1)\}$. Importantly, relaxing beliefs about rationality in such a way does not alter the ranking of IR over DS. Put differently, any such strategic 'top-down' iterative-reasoner should still strictly prefer to play

IR over DS.[15]

Lastly, the comparative statics also hold in Nash equilibrium.[16] IR has a Nash equilibrium in mixed strategies where the equilibrium actions coincide with the actions prescribed by the iterative 'top-down' model of reasoning. The equilibrium payoff is also strictly greater than 12 and strictly dominates the equilibrium payoff in DS, which is exactly 12. The Nash equilibrium of IR is $\big((8/9, 1/9, 0, 0), (0, 13/15, 2/15, 0)\big)$ with payoffs $\big(182/15, 112/9\big)$. DS has a Nash equilibrium in pure strategies: $\big((0, 0, 1), (1, 0, 0)\big)$ with payoffs $(12, 10)$.

The control games The two control games are designed to rule out other confounding factors that can potentially contribute to preferring DS over IR. Their strategic form is depicted in Figure 2. Notice that Player 1's potential payoffs in the two control games are identical to their payoffs in DS, so the only difference between the three games arises from varying Player 2's payoffs.

MS

Player 1 \ Player 2	A	B	C
a	0, 6	12, 10	11, 8
b	5, 16	13, 3	0, 10
c	12, 9	8, 8	0, 10

NE

Player 1 \ Player 2	A	B	C
a	0, 13	12, 14	11, 6
b	5, 5	13, 3	0, 16
c	12, 10	8, 9	0, 9

Figure 2: The controls – The Mixed-Strategy Game (MS) and the Nash-Equilibrium Game (NE)

Our controls serve two purposes. First, we want to control for the size of the game; that is, whether players prefer any smaller game over IR *per se*. To do so, we introduce MS, which is a 3×3 bimatrix game with the iterative 'top-down' model of reasoning prescribing to Player 1 actions $\in \{a, b, c\}$. MS has a Nash equilibrium in mixed strategies similar to IR where players mix over actions $\in \{a, b\}$ (but not c), and Player 1's equilibrium payoff is strictly lower than the

[15]We elaborate on this in Section 5, where we present a more formal analysis.

[16]This is also true in logit Quantal Response Equilibrium. Details are available upon request.

equilibrium payoff in IR.[17]

Second, we want to control for the fact that DS has a unique Nash equilibrium in pure strategies. Thus, we consider NE – a game with a unique Nash equilibrium in pure strategies. In contrast to DS, however, this game is not dominance-solvable. Here too, the iterative 'top-down' model of reasoning prescribes player's action $\in \{a, b, c\}$. Once again, Player 1's equilibrium payoff in NE is strictly lower than the equilibrium payoff in IR. The Nash equilibrium in NE is $\big((0,0,1),(1,0,0)\big)$ with equilibrium payoffs $(12, 10)$, which coincide with the equilibrium payoffs in DS.

As we are solely interested in participants' behavior in the role of Player 1, all three 3×3 games (DS, MS, and NE, respectively) are chosen to share common features. As noted above, all payoffs for Player 1 are kept constant across these games to improve control and ease of comparison. We only altered the payoffs associated with actions $\in \{A, B, C\}$ for Player 2. Moreover, notice that in the control games, like the IR game, all actions are iteratively undominated. Thus, DS stands alone as being the unique game where reasoning about rationality alone is enough to predict the opponent's play.

3 The Experiment

3.1 Implementation

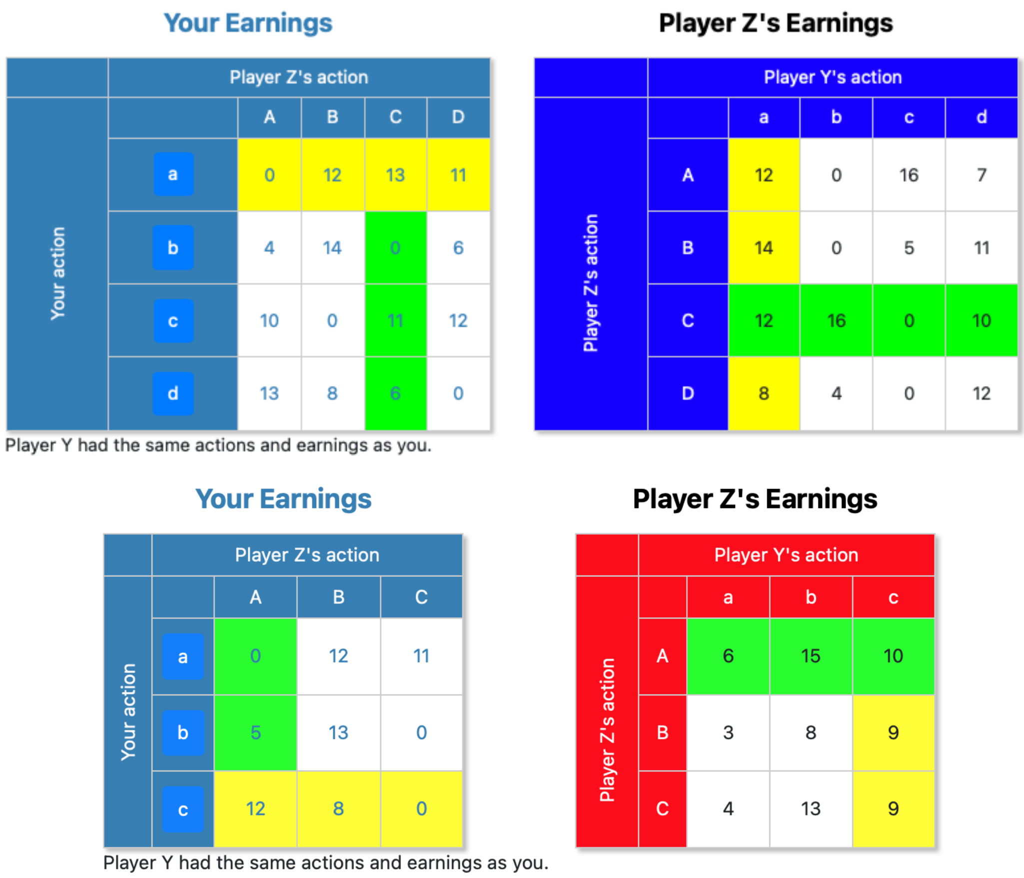
We divided the experiment into two parts. In each part, participants faced four decision-making problems in random order. We told participants that they would be randomly matched with another participant, who had already made their choices in a previous auxiliary session. The purpose of this design feature was to collect all data in an individual decision-making setting, to ameliorate any form of social preferences when choosing actions and participants engaging in forward-induction considerations.

We told participants that this other participant, whom we called "Player Z,"

[17] The Nash equilibrium in MS is $\big((7/9, 2/9, 0),(0, 11/12, 1/12)\big)$ with payoffs $\big(143/12, 76/9\big)$.

is either an undergraduate student from any year or discipline at the University of Toronto or a Ph.D. student in Economics who took several advanced courses that are highly relevant for this experiment. Participants would not learn their opponent type until the end of the experiment. Therefore, participants always made two choices: one if Player Z was an undergraduate student from any year or discipline and another if they were a Ph.D. student in Economics.

Figure 3 visualizes the implementation of the two diagnostic games.



Your Earnings

Your action \ Player Z's action	A	B	C	D
a	0	12	13	11
b	4	14	0	6
c	10	0	11	12
d	13	8	6	0

Player Y had the same actions and earnings as you.

Player Z's Earnings

Player Z's action \ Player Y's action	a	b	c	d
A	12	0	16	7
B	14	0	5	11
C	12	16	0	10
D	8	4	0	12

Your Earnings

Your action \ Player Z's action	A	B	C
a	0	12	11
b	5	13	0
c	12	8	0

Player Y had the same actions and earnings as you.

Player Z's Earnings

Player Z's action \ Player Y's action	a	b	c
A	6	15	10
B	3	8	9
C	4	13	9

Figure 3: Game Implementation – IR (top) and DS (bottom)

The matrices on the left represent participants' payoffs in IR (top) and DS (bottom). The matrices on the right represent Player Z's payoffs in IR and DS, respectively. The opponent type was visualized via color (red = undergraduate and blue = Ph.D. student).

Our experimental implementation of the games makes it particularly salient for participants that Player Z has a strictly dominant strategy in DS. Moreover, in IR, it highlights the attractiveness of action C for the L1 type of Player Z, although it is more nuanced compared to DS. As this type is non-strategic and does not take the other player's incentives into account, visualizing each player's payoffs in a separate matrix directs attention to the sequence of numbers or single entry that is the highest. Put differently, both our design and implementation make natural focal points for a non-strategic player in both games particularly salient.

To improve the experience of participants and to assist in selecting an action, we implemented a highlighting tool that used two colors: yellow and light green. When a participant moved their mouse over a row in their matrix ("Your Earnings"), the action was highlighted in yellow in both matrices: a row in their matrix and a column in Player Z's matrix ("Player Z's Earnings"). By left-clicking the mouse over a row it remained highlighted, and participants could unhighlight it by clicking their mouse again or clicking another row. Similarly, when participants moved their mouse over a row that corresponds to an action of Player Z in "Player Z's Earnings," the row was highlighted in light green, and the corresponding column was highlighted in light green in "Your Earnings." Clicking the mouse over the row kept it highlighted, and clicking it again (or clicking another row) unhighlighted it.

We also told the participants that Player Z participated in a previous auxiliary experimental session in which they were matched with another participant, called "Player Y," who participated in the same session and played their role. When Player Z was an undergraduate student in any year or discipline, so was Player Y; and when Player Z was a Ph.D. student in Economics, so was Player Y. We used Player Z's decisions from the auxiliary sessions to determine participants' earnings in the main experiment.

In addition, we gave participants the opportunity to write notes to their "future self." Below each decision problem, participants could write down the reasoning behind their choice of action in a text box. What they typed was displayed later on in the experiment. We told participants that these notes would help them make decisions in the second part of the experiment.

To account for possible order effects, we gave participants another opportunity to revisit their choices and confirm them.[18] We displayed their notes and participants were able to modify them. Afterwards, participants advanced to the next part of the experiment.

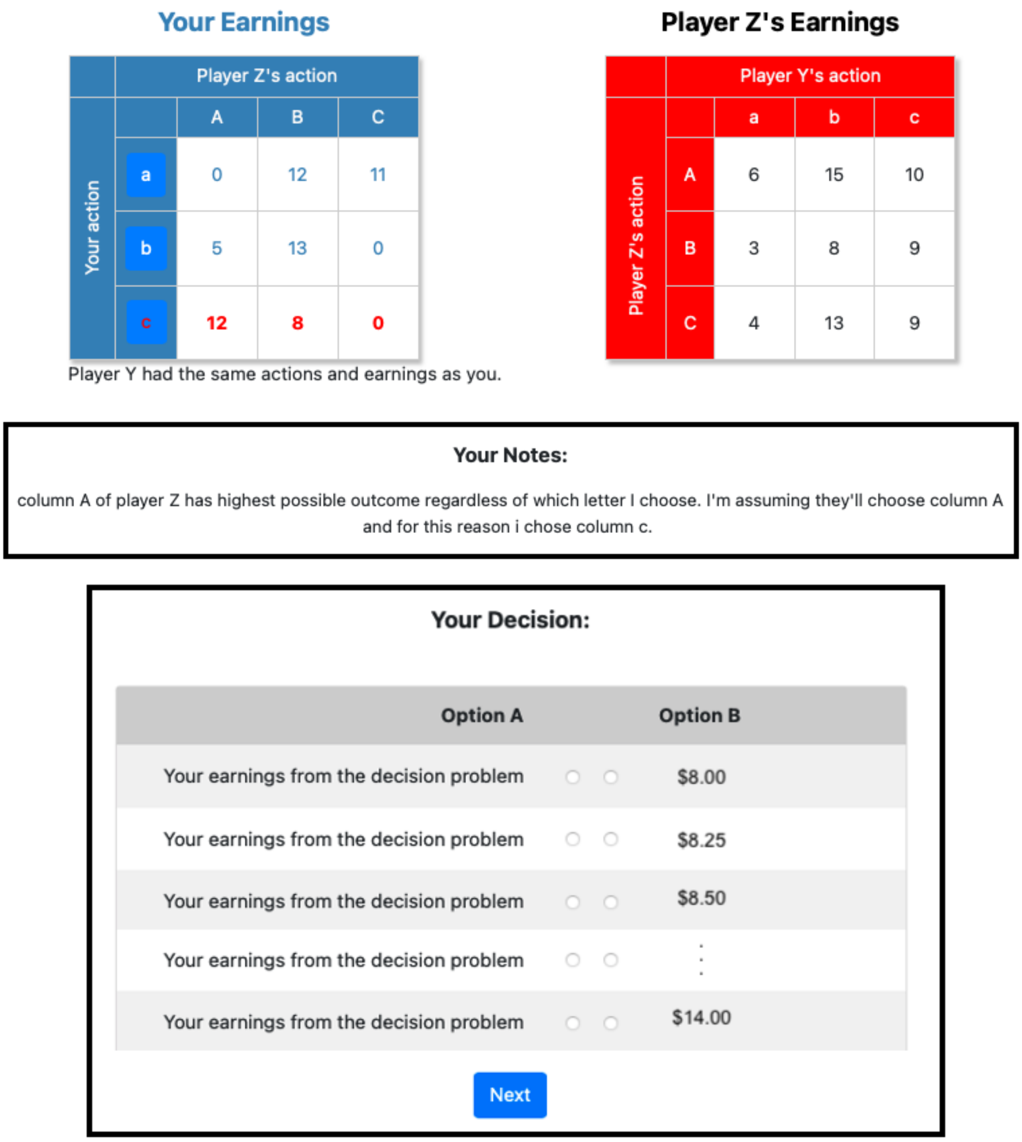


Figure 4: The Valuation Task

In the second part of the experiment, we elicited participants' approximate valuations via choice lists. We asked them to make a series of choices between

[18]We find no evidence of order effects, using both parametric and non-parametric tests.

playing the four decision problems against both Player Z types with their action choices from the first part of the experiment and sure amounts. For example, suppose that in the first part of the experiment a participant chose action c in any given 3×3 game, as highlighted in Figure 4. The payoff from the decision problem depends on the action chosen by Player Z and is either \$12, \$8, or \$0 if Player Z chose A, B, or C, respectively.

The choice problems were organized in four pairs ($4 \times 2 = 8$ lists), where Option A changed between lists and represented participants' payoffs from each of the four decision problems against both types of opponent from the first part of the experiment. Option B paid with certainty and started at \$8 in the decision of the choice list, and increased by \$0.25 as the participant moved from one line to the next until \$14. For each decision problem, we showed participants their notes from the first part of the experiment to remind them of their reasoning behind their action choices.

Finally, one of the choice problems in one of the choice lists was randomly selected, and the participant's choice in that choice problem determined their payment. If a participant chose the sure amount in Option B, then they received the payment specified in Option B in that choice problem. If a participant opted for Option A, then their payment depended on the action chosen in the decision problem in the first part of the experiment, if their Player Z was an undergraduate student or a Ph.D. student, and on the action chosen by Player Z.[19]

3.2 Participants and Procedure

We conducted the experiment in April 2020 with students enrolled at the University of Toronto. Participants were recruited from the Toronto Experimental Economics Laboratory's (TEEL) subject pool using ORSEE (Greiner 2015). No one participated in more than one session. Participants signed up ahead of time for a particular day, either the 4^{th} or 5^{th} of April 2020 for the auxiliary part of the experiment; or the 11^{th}, 13^{th}, and 15^{th} to 20^{th} of April 2020 for the main

[19] The timeline of the experiment and the key features are visualized in the Supplemental Appendix.

experiment. On the day of the experiment, we sent an electronic link to participants at 8 AM EDT, and they had to complete the tasks by 8 PM EDT. During this time window, participants could contact an experimenter anytime via cell phone or Skype for assistance. After reading the instructions, participants had to correctly answer nine incentivized comprehension questions before starting the first task, and five more incentivized comprehension questions before starting the second task. We paid \$0.25 for answering each question correctly on their first attempt. If participants made a mistake, no payment was made for that question, but they had to answer it correctly to proceed to the next question. The experiment was programmed in oTree (Chen, Schonger, and Wickens 2016). We recruited a total of 244 participants (9 for the auxiliary sessions and 235 for the main experiment) and all payments were made via Interac e-transfer, a commonly used payment method by Canadian banks that only requires an e-mail address and a bank account. The average participant earned approximately \$18 (maximum payment was \$22.50 and minimum payment was \$5.50), including a show-up payment of \$5. All payments were made in Canadian dollars. The instructions and experimental interface are reproduced in the Supplemental Appendix.[20]

3.3 Discussion of the Implementation and Procedure

The core idea of this paper is to identify a novel behavior that reflects whether reasoning is outside the iterative 'top-down' model of reasoning. Thus far, we have developed an identification strategy for such behavior and, before presenting the results of the evaluation of its pervasiveness, we briefly discuss some aspects of the experimental implementation and its procedure. We collected Player Z's decisions on action choices in the four games in two separate auxiliary sessions. This has the following advantages: First, we were able to match participants (Player Y and Player Z) with the same level of sophistication. Second, we could collect all the decisions in the main experiment in an "individual decision-making" framework. As we collected data during the lockdown in the COVID-19 pandemic, we could not

[20] A live version with all dynamic elements displayed to participants can be accessed upon request.

run any experiment sessions in the laboratory. Instead, undergraduate students enrolled at the University of Toronto eagerly participated remotely. Thus, we were able to avoid any coordination issues stemming from simultaneous strategic decision-making in an online context. Lastly, as choices and payments in the auxiliary sessions had materialized already, this design can eliminate concerns that choices made by participants in the main experiment were motivated by social preferences or forward induction considerations. To avoid quick heuristic-based decision-making, we forced participants to spend at least 10 minutes on each set of instructions and at least 3 minutes on each of the four games against either opponent type before buttons were activated. Further, we presented all four games in random order to avoid any order effects and, in addition, gave participants the opportunity to revise their decisions after they were exposed to all four games and had selected an action choice. Remaining conscious of possible order effects, we also reversed the opponent order between the two parts of the experiment. That is, if participants faced always an undergraduate student before a Ph.D. student in Economics when choosing an action, then they always faced a Ph.D. student in Economics before an undergraduate student in the valuation task and *vice versa*.

4 Results

We break the analysis into four sections. After a brief coherence examination of the valuation data, we begin our main analysis by presenting the experimental results, focusing first on the preferences between IR and DS, and then explore the valuation data in all four games. Next, we focus on behavior conditional on the opponent's identity; that is, whether Player Z was an undergraduate student of any year or discipline or a Ph.D. student in Economics. Lastly, we delve into the non-choice data embedded in the participants' notes.

4.1 Coherence of Elicited Valuations

Before turning to choice behavior and the ranking of IR and DS, we first present the empirical valuation data from some of the games to illustrate both that participants exhibit reasonable valuations and that there are powerful insights to be gained for an outside observer by eliciting participants' certainty equivalent for each game.

In total, we collected data from $N = 235$ participants. The only exclusion restriction for valuations that we impose is *consistency with rationality*. That is, we exclude behavior characterized by valuations that exceed the maximum possible payoff given their action choice; for example, playing action b with a valuation $v = 14$ in DS, MS, or NE, respectively. Figure 5 displays several empirical value distributions.

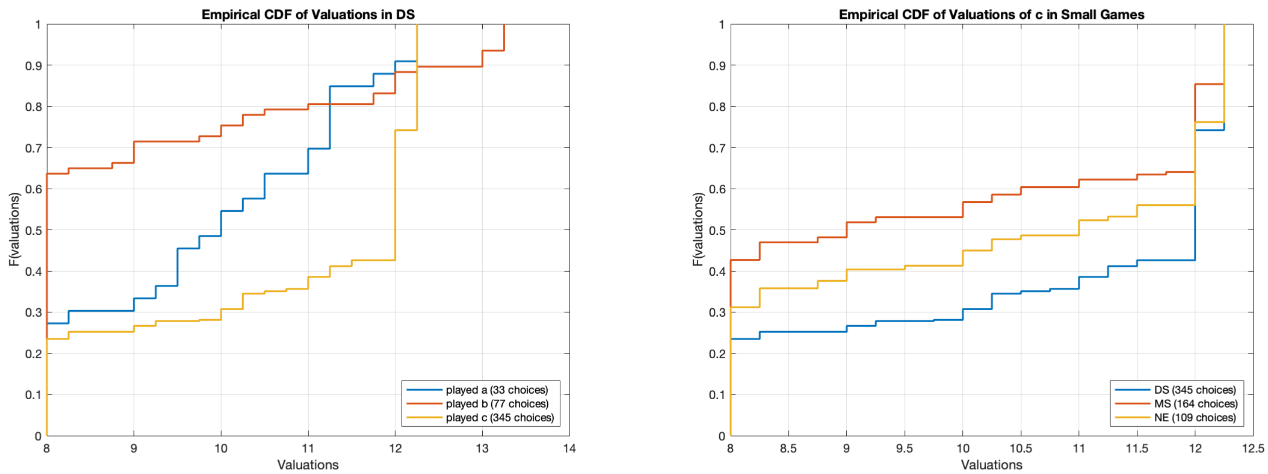


Figure 5: Empirical Value Distributions of DS by action choice; and Empirical Value Distributions of DS MS, and NE conditional on Playing Action c in DS

First, we show the empirical value distributions in DS by action for $n = 455$ choices; that is, all choices with consistent valuations irrespective of opponent type. Approximately 76% of choices fall on action c, 17% play action b and the remaining 7% choose action a. Participants who play c tend to value playing DS more than participants who choose a or b. Recall that c is dominated by a according to $\succ_1$, and that the highest payoff in b (13) is higher than the highest payoff in c (12). This suggests that those who played c have done so for strategic reasons.

Second, we highlight the empirical value distributions in DS and both control games conditional on playing action c irrespective of opponent type, leaving us with $n = 618$ choices in total. Recall that participants face the exact same payoffs in these three games, so different choices and valuations in these games must arise from the different strategic structures. The frequency of action-c play in DS is approximately $2-3$ times higher compared to those in the two control games, MS and NE, respectively. Furthermore, the empirical value distribution for DS first-order stochastically dominates those for NE and MS, suggesting that opponent behavior in DS is easier to predict relative to NE and MS.

4.2 *IR* and *DS* Valuations

We impose an additional exclusion restriction for the IR and DS choices in our main analysis. That is, in addition to imposing consistency of rationality, we focus on observed choices where only action c is played in DS. Restricting attention to action c in DS allows us to isolate the choices made by strategic participants, as the L1 non-strategic type only plays actions a or b in DS and never plays action c. Thus, we restrict attention to $n = 343$ choices. That is, we focus on 179 participants facing an undergraduate student and 164 participants facing a Ph.D. student in Economics.[21] To give a first overview, we present the aggregate results of the action choices in the diagnostic games. Table 1 offers a synopsis of the frequency of action choices in IR.

Table 1: Frequency of Action Choices in the IR Game

Action	IR
a	242/343
b	37/343
c	37/343
d	27/343

All choices made irrespective of opponent type.

[21] All analyses reported in the main text are replicated for all participants and choices in our sample. These results are reported in the Supplemental Appendix.

Approximately 71% of the choices in IR are concentrated on action a, and the remainder is roughly equally distributed among actions b, c, and d, respectively.

As a first-pass, we summarize the choice behavior and the ranking of IR and DS irrespective of the opponent type. Table 2 lists these results.

Table 2: Preferences between IR and DS

	$IR \succ DS$	$IR \precsim DS$
IRM Prediction	*all*	*nil*
Ratio	154/343	189/343
Percentage	44.9%	55.1%

All choices made irrespective of opponent type.
$IRM \equiv$ Iterative 'top-down' model of reasoning.

The observed choices are clearly at odds with the predictions of the iterative 'top-down' model of reasoning (or Nash equilibrium). While players are predicted to strictly prefer IR over DS, less than half of all observed choices are in line with the prediction. This is the first evidence at the aggregate choice-level suggesting that participants' reasoning may fall outside the iterative 'top-down' model of reasoning.

Introducing controls As a next step, we include the two control games in our aggregate-choice analysis. We are interested in those participants who weakly prefer DS over IR, and not those who may have a preference for smaller games or Nash equilibrium in pure strategies *per se*.

To do so, we require that participants make choices consistent with best-responding in both MS and NE games.[22] As a result, we are now focussing on 153 participants facing an undergraduate student and 138 participants facing a Ph.D. student in Economics, respectively. Table 3, Control #1 lists these results of $n = 291$ observed choices irrespective of opponent type. As is evident, controlling for best-response consistency at the aggregate choice level does not make a substantial dent in participants' overall ranking of IR and DS.

[22]In this step, we remove participants' choices of a with a valuation $v > 12$, and further exclude those whose valuations that exceed the maximum possible payoff given their action choice in either of the two control games.

Table 3: Controlling for Best-Response Consistency in All Games
and Equal Valuations of All Small Games

	$IR \succ DS$	$IR \precsim DS$
IRM Prediction	*all*	*nil*
Control #1 B-R Consistency	135/291 46.4%	156/291 53.6%
Control #2 NE Preference	138/268 51.5%	130/268 48.5%
Control #3 Equal Valuations	107/213 50.2%	106/213 49.8%

All choices made irrespective of opponent type excluding all choices that are inconsistent with best-responding ("C#1"); preference for Nash equilibrium in pure strategies ("C#2"); and value DS, MS, and NE equally ("C#3"). $IRM \equiv$ Iterative 'top-down' model of reasoning.

Next, we exploit the Nash equilibrium in pure strategies that characterizes both DS and NE. Here, we exclude action c choices in both games *and* value NE weakly above IR. This allows us to control for those that may feature an intrinsic preference for Nash equilibrium in pure strategies *per se*. By doing so, we focus on 147 participants playing against an undergraduate student and 78 participants playing against a Ph.D. student in Economics, respectively. The summary statistics for this control are listed as Control #2 in Table 3. Similar to the previous control, this control does not alter the overall ranking of the diagnostic games either.

Finally, we leverage MS and NE and, in this step, exclude those choices that value all small games equally; i.e., $v_{DS} = v_{MS} = v_{NE}$. This allows us to control for those participants who have high valuations in DS relative to IR due to an intrinsic preference for smaller games or Nash equilibrium in pure strategies. This results in concentrating on 113 participants playing against an undergraduate student and 100 participants playing against a Ph.D. student in Economics. These results are reported in Table 3, Control #3. Although this control slightly reduces the proportion of choices that rank the DS game higher than the IR game, about half of the choice and valuation decisions are inconsistent with the iterative reasoning model. In general, the inclusion of the controls does not alter the results. Although the ratio of those who weakly prefer DS over IR somewhat decreases, the big picture still suggests that the reasoning of many participants may fall outside of the iterative 'top-down' model.[23]

Belief that opponent is rational Here, we consider those participants who believe that their opponents are rational and are confident that Player Z is rational. Recall that our design makes the sharpest predictions for these types – unambiguously predicting that participants using the iterative 'top-down' model of reasoning would strictly prefer to play IR over DS. Our design allows us to identify these participants by exploiting the valuation data collected in the second part of our experiment. In particular, we now include an additional exclusion restriction by requiring valuations of $12 \leq v \leq 12.25$ in DS.[24] Table 4 summarizes the choice behavior by the ranking of IR and DS irrespective of the opponent type but conditional on believing in the opponent's rationality.

Table 4: Belief that Opponent Is Rational

	$IR \succ DS$	$IR \precsim DS$
IRM Prediction	*all*	*nil*
Ratio	72/197	125/197
Percentage	36.5%	63.5%

All choices made irrespective of opponent type conditional on believing in opponent's rationality.
$IRM \equiv$ Iterative 'top-down' model of reasoning.

When requiring players to be confident that their opponent is rational (the value of c in DS is 12, indicating that Player 1 is confident that Player 2 will play the dominant action), close to two-thirds of $n = 197$ choices rank DS above IR. This behavior reflects reasoning that falls outside the iterative 'top-down' model.

[23] A potential concern may arise because we used choice lists to elicit participants' approximate valuation for each game. As these lists are discrete, we could potentially misclassify participants' ranking. Those participants who valued both IR and DS *exactly* at $v = 12.25$ could be classified as weakly ranking DS above IR despite being consistent with the iterative 'top-down' model of reasoning. Of the $n = 343$ choices presented in Table 2, only 29 choices value both games exactly at $v = 12.25$. For the controls, this reduces further to $^{10}/_{291}$ in Control #1 and $^{7}/_{213}$ in Control #3, respectively.

[24] This results in concentrating on 106 (91) participants playing against an undergraduate student (a Ph.D. student in Economics).

4.3 Opponent Type

We now turn to choices at the subject-level and discuss differences in behavior by opponent type. We maintain all our exclusion restrictions discussed above, but as we are interested in participants that satisfy these exclusion restrictions against *both* opponent types – the intersection – we thus concentrate now on $n = 144$ participants. Thus far, we have established that approximately half of the choices fall outside the iterative 'top-down' model of reasoning. Recall that this turns out to be true especially if one believes that their opponents are rational. Among this subset of choices, approximately two-thirds of choices fall outside the model.

Table 5 shows the comparative statics of the ranking over the set of diagnostic games conditional on the opponent's identity; that is, whether participants played against an undergraduate student of any year or discipline or a Ph.D. student in Economics.

Table 5: Ranking of IR and DS by Opponent Type

			Undergraduate $IR \succ DS$	*Undergraduate* $IR \precsim DS$
Ph.D.	$IR \succ DS$	IRM Prediction	*all*	*nil*
		Ratio	46/144	23/144
		Percentage	31.9%	16.0%
	$IR \precsim DS$	IRM Prediction	*nil*	*nil*
		Ratio	18/144	57/144
		Percentage	12.5%	39.6%

$IRM \equiv$ Iterative 'top-down' model of reasoning.

These numbers are not overly sensitive to the opponent's type: 71.5% of participants exhibit a stable model of reasoning irrespective of the opponent's characteristics. That is, the majority of participants respond similarly to both undergraduate and Ph.D. students in Economics. Specifically, about 32% of the participants' choices are consistent with the iterative 'top-down' model of reasoning against both undergraduate students and Ph.D. students in Economics in IR and about 40% are inconsistent against both.[25] Among the remainder, of those

[25]These 57 participants value the DS game (weakly) more than the IR game. Moreover, the

who respond to the opponent's type, the results are split. 12.5% are consistent with the iterative 'top-down' model of reasoning against undergraduate students and not Ph.D. students in Economics, while 16% are consistent with the iterative reasoning model against Ph.D. students in Economics but not undergraduate students.

4.4 Non-Choice Data

Recall that we gave participants the opportunity to write notes to their "future-self." Below each of the two diagnostic games and the two control games against either opponent type, participants could write down the reasoning behind their choice of action in a text box. If participants decided to type anything in these text boxes, it was displayed again later in the experiment: the first time when participants were prompted to confirm their choice of action and the second time when faced with the valuation task. We did not force participants to write anything in these text boxes; however, we told them that these notes would help them when making choices in the second part of the experiment. As expected, not all participants made use of this opportunity. However, those who did give us the opportunity to use their notes as "the window of the strategic soul."[26] Using both action choice and valuation data, we documented evidence at the aggregate choice-level that suggests that participants may value the predictability of their opponents' behavior. Moreover, we showed that this observation is even more stark if participants believe that their opponents are rational with 63.5% of choices ranking DS above IR. Among this subset of participants, we are curious to see whether there is any suggestive evidence of participants indicating that the opponents' actions are predictable in DS and IR, and if there is any difference by the ranking of IR and DS. We have established that 197 choices are consistent with holding the belief that their opponent is rational, meaning that the player is confident that Player Z is rational. In 105 (113) of these choices, participants

valuation data reveal that for these participants the IR game becomes relatively more valuable than the DS game when playing against a Ph.D. student rather than an undergraduate student.

[26] Vincent Crawford coined this term in Crawford (2008).

took notes in IR (DS). Table 6 provides summary statistics for this subset of choices by the ranking of the set of diagnostic games.

Table 6: Notes – Belief that Opponent Is Rational

Indication that Player Z's Action Is Predictable

		IR		DS	
		yes	*no*	*yes*	*no*
$IR \succ DS$	Ratio	22/52	14/53	18/60	25/53
	Percentage	42.3%	26.4%	30.0%	47.2%
$IR \precsim DS$	Ratio	30/52	39/53	42/60	28/53
	Percentage	57.7%	73.6%	70.0%	52.8%

If a participant indicated that the opponent's choice is predictable in one of the games, it increased the likelihood that they would prefer that game. For example, out of the 105 participants who took notes in the IR game, 52 participants noted that Player Z's action is predictable. The probability of preferring IR to DS increased from 26.4% to 42.3% (an increase of approximately 60%). Similarly, of the 113 participants who took notes in DS, 60 wrote that Player Z's action was predictable. The probability of preferring the IR game to the DS game among them was 30%, compared to 47.2% among participants who took notes but did not mention the predictability of Player Z's action in DS (a decrease of more than 36.4%).

We complement this qualitative analysis with natural language processing tools to gain additional insights on participants' thought process. In line with the choice behavior presented in Section 4, participants who rank one diagnostic game above the other also express their reasoning in more detail, use more complexity-related keywords to express more sophisticated reasoning, are more positive and optimistic, and feature more determination and certainty in their preferred game compared to the other diagnostic game. Moreover, differences in the ranking of games are associated with different topics and clusters that can be recovered using natural language models.[27] This lends further qualitative support to the idea that the $DS \succsim IR$ group and the $DS \prec IR$ group treat the two diagnostic games

[27]We elaborate on this in detail in the Supplemental Appendix.

systematically differently and employ fundamentally different reasoning processes.

5 Theoretical Analysis

In Section 2, we provided intuitive explanations for our identification strategy. In this section, we elaborate and present a formal analysis.

5.1 Theory

Let $G = (S_1, S_2, u_1, u_2)$ be a finite 2-player game where S_i is player i's strategy set and $\pi_i : S_1 \times S_2 \to \mathbb{R}$ is player i's pecuniary payoff function, which depends on player i and the other player's $(-i)$ strategies. We allow for general expected-utility preferences over monetary payoffs. Let $\mathcal{U}$ be the set of von Neumann-Morgenstern utility functions, which are strictly increasing functions mapping $\mathbb{R}$ to $\mathbb{R}$. For any $u_i \in \mathcal{U}$, the function $u_i \circ \pi_i : S_i \times S_{-i} \to \mathbb{R}$ represents the utility of player i. Denote by $\mu_{-i} \in \Delta(S_{-i})$ player i's beliefs over player $-i$'s strategies. Extend $u_i\big(\pi_i(S_i, S_{-i})\big)$ to $u_i\big(\pi_i(S_i, \mu_{-i})\big)$ in the usual way to represent player i's expected utility.

Let $\mathbb{BR}_i$ be the best response set for each player i. This set specifies the strategies that are a best response for player i given both player i's preferences, $u_i \in \mathcal{U}$, and the belief they hold about the play of the other player, μ_{-i}. Formally, for $u_i \in \mathcal{U}$ and $\mu_{-i} \in \Delta(S_{-i})$,

$$\mathbb{BR}_i[u_i, \mu_{-i}] := \{s_i \in S_i : u_i\big(\pi_i(s_i, \mu_{-i})\big) \geq u_i\big(\pi_i(r_i, \mu_{-i})\big), \text{ for each } r_i \in S_i\}.$$

We will be interested in two solution concepts. First, the iterative 'top-down' model of reasoning, which intuitively captures how players reason when they can model the behavior of others. Second, the concept of 2-rationalizable strategies, which incorporates the assumption that player i is rational and believes player $-i$ is rational and nothing more. Intuitively, this solution concept captures how players reason when they cannot model the behavior of others. We define both below.

Iterative 'top-down' model of reasoning This model is anchored by the non-strategic L1 behavior characterized by $\succ_1$. Let $L_i^1 = \{s_i \in S_i | \nexists\ s_{i'} \in S_i \text{ where } s_{i'} \succ_1 s_i\}$ be the set of actions that can be played by the L1 type. This is the set of actions that are undominated according to $\succ_1$.

In Section 2, we discussed the possibility of extending the model to allow for uncertainty over others' rationality. We do this by defining an L0 type that is non-strategic and plays all actions – even strictly dominated actions – with positive probability. Specifically, we impose the restriction that the L0 type plays uniformly random: $\mu_i^0(s) = \frac{1}{|S_i|}$ for all $s \in S_i$. Strategic types that place positive probability on facing the L0 type will be uncertain about the rational play of others.

The behavior of all Lk types can be defined recursively, anchored on the behavior of the L0 and L1 types. Denote by L_i^k the set of actions consistent with k iterations of reasoning by player i. Then, for $k \geq 2$, the set L_i^k is the set of strategies s_i in $\mathbb{BR}_i[u_i, \mu_{-i}]$ such that there exists some $u_i \in \mathfrak{U}$ and $\mu_{-i} \in \Delta(S_{-i})$ that satisfies the following two conditions. First, beliefs over the play of others must take the following form: $\mu_{-i} = p \cdot \mu_{-i}^0 + (1-p) \cdot \eta_{-i}$ for some $p \in [0, 1)$ and $\eta_{-i} \in \Delta(S_{-i})$ with $\eta_{-i}(\cup_{j=1}^{k-1} L_{-i}^j) = 1$. This ensures that player i's beliefs about player $-i$'s behavior are consistent with the assumption that players' reasoning is organized in a 'top-down' fashion. Put differently, player i can only assign positive probability to actions played by types with levels strictly less than k. Second, $s_i \in \mathbb{BR}_i[u_i, \mu_{-i}]$. This condition ensures that player i's strategy s_i maximizes their expected utility given player i's preferences u_i, and the belief that player $-i$ plays according to μ_{-i}. We will refer to any action a_i in L_i^k as *an action played by the Lk type* for player i.

2-*rationalizability* The solution concept of 2-rationalizable strategies incorporates both the assumption of rationality and belief in rationality. Let S_i^1 be the set of strategies s_i such that there exists some $u_i \in \mathfrak{U}$ and $\mu_{-i} \in \Delta(S_{-i})$ with $s_i \in \mathbb{BR}_i[u_i, \mu_{-i}]$. The set S_i^1 includes all rational strategies for player i. These are a best response for player i given their preference u_i and beliefs μ_{-i} about player $-i$'s play. We refer to any action a_i in S_i^1 as a 1-*rationalizable strategy.* Given this, we can define S_i^2 as the set of strategies s_i so that there exists some $u_i \in \mathcal{U}$ and $\mu_{-i} \in \Delta(S_{-i})$ that satisfies the following conditions. First, $s_i \in \mathbb{BR}_i[u_i, \mu_{-i}]$, which ensures that s_i maximizes player i's expected utility given the belief that player $-i$ behaves according to μ_{-i}. Second, $\mu_{-i}(S_{-i}^1) = 1$. This ensures that player i can only place positive probability on 1-rationalizable strategies, which are the strategies consistent with the assumption that player $-i$ is rational. We will refer to any action s_i in S_i^2 as a 2-*rationalizable strategy.*[28]

5.2 Revisiting the Diagnostic Games

The iterative-reasoning game "IR" First, note that we can denote any probability measure $p \in \Delta(S_1)$ (and $p \in \Delta(S_2)$, respectively) as a 4-tuple (p_1, p_2, p_3, p_4). This represents the probabilities over $\{a, b, c, d\}$ (and $\{A, B, C, D\}$, respectively). Then in this game, L0 behavior is given by $\mu^0 = (1/4, 1/4, 1/4, 1/4)$ for both players. Further, recall from Section 2 that $L_1^1 = \{a, b\}$ and $L_2^1 = \{C\}$.

The L_i^k sets can then be calculated recursively given the anchoring L0 and L1 behavior. Let $k \geq 2$. For Player 1, the Lk type can hold any belief about Player 2's behavior that is a mixture between μ^0 and the two degenerate beliefs: $(0, 1, 0, 0)$ and $(0, 0, 1, 0)$. In other words, beliefs take the form $\mu_2 = \left(p_0/4, p_0/4 + p_B, p_0/4 + p_C, p_0/4\right)$ for some $p_0, p_B, p_C \in [0, 1]$ with $p_0 + p_B + p_C = 1$. A strategy s_i is in L_1^k if there exists some $u \in \mathcal{U}$ such that $s_i \in \mathbb{BR}_1[u, \mu_2]$. Clearly, actions a and b are in L_1^k as they maximize the expected payoff under the player's belief when $p_C = 1$ and $p_B = 1$, respectively. Importantly, we also need to ensure that a and b are the only choices that maximize expected utility for every utility function u.[29] We begin with the observation that a strategy $s_i \in S_1$ induces a lottery through the belief $p \in \Delta(S_2)$, which we denote $s_{i,p}$. For example, the action a induces the

[28] In order for the solution concept to be free of assumptions about risk preferences we explicitly allow players to hold any expected utility preferences. The same result could be achieved by specifying a single preference specification for each player with preferences characterized by extreme risk aversion. This follows from Battigalli, Cerreia-Vioglio, Maccheroni, and Marinacci (2016) and Weinstein (2016) who show that risk aversion expands the set of k-rationalizable actions (while risk loving contracts the set).

[29] For this we will rely on the following equivalence: a lottery p first-order stochastically dominates lottery q if and only if p is preferred to q for all $u \in \mathcal{U}$.

lottery $a_{\mu_2} = (13, {p_0}/{4}; 12, {p_0}/{4} + p_B; 11, {p_0}/{4} + p_C; 0, {p_0}/{4})$. This lottery first-order stochastically dominates the lotteries c_{μ_2} and d_{μ_2}. It follows that actions c and d cannot maximize the player's expected utility for any utility function u. Thus, we conclude that $L_1^k = \{a, b\}$.

For Player 2, the Lk type can hold any belief about Player 1's behavior that is a mixture between μ^0 and the two degenerate beliefs: $(1, 0, 0, 0)$ and $(0, 1, 0, 0)$. In other words, beliefs take the form $\mu_1 = \left({p_0}/{4} + p_a, {p_0}/{4} + p_b, {p_0}/{4}, {p_0}/{4}\right)$ for some $p_0, p_a, p_b \in [0, 1]$ with $p_0 + p_a + p_b = 1$ and $p_0 < 1$. Consider the case where $p_a \neq 1$, then the lottery C_{μ_1} first-order stochastically dominates the lotteries A_{μ_1} and D_{μ_1}. Next, consider the case where $p_a = 1$, then the lottery B_{μ_1} first-order stochastically dominates the lottery x_{μ_1} for $x \in \{A, C, D\}$. Thus, we conclude that $L_2^k = \{B, C\}$.

$$L_1^k = \{a, b\} \text{ if } k \geq 1 \qquad\qquad L_2^k = \begin{cases} \{C\} & \text{if } k = 1 \\ \{B, C\} & \text{if } k \geq 2 \end{cases}$$

We now turn to characterizing the 2-rationalizable set for Player 1, which captures the case of a player who is rational and believes that Player 2 is rational. Here, Player 1 believes that Player 2 plays a 1-rationalizable strategy. The 2-rationalizable set for Player 1 and the 1-rationalizable set for Player 2 are:

$$S_1^2 = \{a, b, c, d\} \qquad\qquad S_2^1 = \{A, B, C, D\}$$

It is straightforward to see that all actions for Player 2 are 1-rationalizable. This is the case as each action maximizes expected payoffs under some degenerate belief about the play of Player 1. It follows that all actions are 2-rationalizable for Player 1 as each action for Player 1 maximizes expected payoffs under some degenerate belief about Player 2's behavior.

Lastly, we elicited participants' valuation for each game, i.e., their certainty equivalent. Since a player's utility function is monotone, the analyst can infer their ranking over the games. Moreover, the valuations reveal important information about participants' beliefs.

In the iterative 'top-down' model of reasoning, restricting attention to types that are rational and believe that their opponent is rational confines attention to types that assign zero weight on others being the L0 type. The expected payoff in IR must be *strictly greater than 12* for these types. It is straightforward to confirm this claim by setting $p_0 = 0$ in the above arguments. This means that any type holds a belief that is a mixture of $(0, 1, 0, 0)$ and $(0, 0, 1, 0)$. For any such belief $\mu_2 = p(0, 1, 0, 0) + (1 - p)(0, 0, 1, 0)$, the certainty equivalent of the lottery $a_{\mu_2} = \big(12, p; 13, (1 - p)\big)$ is above 12 whenever $p \neq 1$, and the certainty equivalent of the lottery $b_{\mu_2} = \big(14, p; 0(1 - p)\big)$ is 14 whenever $p = 1$. To summarize, players who are rational and hold the belief that their opponents are rational believe that they can guarantee themselves a payoff that is strictly greater than 12. It follows that the certainty equivalent of IR for any expected utility player who believes that their opponent is rational is strictly above 12.

Caution is potentially warranted if Player 1 is ambiguity averse as they may value IR at 12. This, however, can only occur under an extreme form of ambiguity aversion coupled with the player holding degenerate beliefs. More precisely, it requires Player 1 to play the "safe" action a, to have maxmin expected-utility preferences *and* their set of priors must include beliefs that Player 2 plays B with certainty and a prior that assigns a probability strictly less than $6/7$ that Player 2 plays B.[30]

Moving to payoffs when applying the concept of 2-rationalizability. A player that believes others are rational can hold any belief over Player 2 choosing a 1-rationalizable action. This means that in IR Player 1 can hold any belief about the play of Player 2. In this case, such players may *not* believe that they can guarantee themselves any certain payoff. Moreover, one might reasonably conjecture the

[30]Whether this is an important concern is an empirical question. We can exploit participants' actions and valuations in the control games to evaluate if ambiguity aversion governs participants' valuations. If we allow for maxmin expected utility preferences, and allow that the set of priors of a player of level $(k+1)$ includes all degenerate priors consistent with the strategies in L_2^k in the control games, then (for any action in) both MS and NE have to be valued at 8. In our data, of all choices, only 1 choice exhibits such extreme form of ambiguity aversion.

certainty equivalents of these actions to be less than 12.

The dominance-solvable game "DS" In this game, the L0 behavior is given by the 3-tuple $\mu^0 = (1/3, 1/3, 1/3)$ for both players. Further, recall from Section 2 that $L_1^1 = \{a, b\}$ and $L_2^1 = \{A\}$.

The L_i^k sets can then be calculated recursively given the anchoring L0 and L1 behavior. Let $k \geq 2$. For Player 1, the Lk type can hold any belief about Player 2's behavior that is a mixture between μ^0 and the degenerate belief: $(1, 0, 0)$. In other words, beliefs take the form $\mu_2 = \left(p_0/3 + p_A, p_0/3, p_0/3\right)$ for some $p_0, p_A \in [0, 1]$ with $p_0 + p_A = 1$. A strategy s_i is in L_i^k if there exists some $u_i \in \mathcal{U}$ such that $s_i \in \mathbb{BR}_i[u_i, \mu_{-i}]$. Clearly, action a and c are in L_1^k as they maximizes the expected payoff under the player's belief when $p_0 = 1$ and $p_A = 1$ respectively. Further, notice that the lottery b_{μ_2} is not first-order stochastically dominated by either lotteries a_{μ_2} or c_{μ_2}, this means we can find some $u_i \in \mathcal{U}$ such that $b \in \mathbb{BR}_1[u_i, \mu_2]$. Thus, $L_1^k = \{a, b, c\}$.

Turning to the behavior of the Lk type of Player 2, this type can hold any belief about Player 1's behavior that is a mixture between μ^0 and the degenerate beliefs: $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$. In other words, a Lk type can hold any beliefs over Player 1's play, $\mu_1 \in \Delta(S_1)$. Notice, however, that Player 2 has a strictly dominant strategy, this means that A is always the best response for Player 2 regardless of her beliefs. In other words, the lottery A_{μ_1} first-order stochastically dominates the lotteries B_{μ_1} and C_{μ_1}. Thus, we conclude that $L_1^k = \{A\}$.

$$L_1^k = \begin{cases} \{a, b\} & \text{if } k = 1 \\ \{a, b, c\} & \text{if } k \geq 1 \end{cases} \qquad\qquad L_2^k = \{A\} \text{ if } k \geq 1$$

Lastly, we briefly discuss the 2-rationalizable predictions. Again, since A is strictly dominant for Player 2, it is the unique 1-rationalizable action. It follows that the only 2-rationalizable action for Player 1 is c.

$$S_1^2 = \{c\} \qquad\qquad S_2^1 = \{A\}$$

In this game, a rational type who believes that their opponent is rational must hold beliefs of the form $(1, 0, 0)$. Such players believe that they can guarantee themselves a payoff of *exactly 12* with certainty. Notice that reasoners who cannot model Player 2's behavior – beyond the belief that Player 2 should play a 1-rationalizable strategy – might reasonably rank DS above IR.

If Player 1 plays c and values the game less than 12 it reveals to the analyst that the player is not confident that Player 2 is rational, since the certainty equivalent of the lottery induced by c is lower than 12 only if it assigns a strictly positive probability that Player 2 will choose a dominated action. Further, such valuations shed light on whether the simpler iterative reasoning model from Section 2 or the more general iterative 'top-down' model of reasoning that explicitly allows for uncertainty over rationality predicts participants' behavior more accurately.

Player 1's preferences over IR and DS We first restrict our attention to players that are rational *and* believe that their opponents are rational. Consider the preferences of such types over the two diagnostic games: IR and DS. Although DS has a smaller strategy space compared to IR and is dominance-solvable, the game's expected payoff of 12 is strictly lower than the expected payoff of IR in the iterative 'top-down' model of reasoning. In other words, a 'top-down' iterative-reasoner should strictly prefer to play IR over DS. We now relax the assumption of belief in rationality. When considering the iterative 'top-down' model of reasoning, this means that we allow players to place positive weight on the L0 type. Fix $p_0 \in [0, 1)$ as the probability assigned to the L0 type. In IR, the belief of a 'top-down' reasoner takes the following form: $\mu_2^{IR} = p_0(1/4, 1/4, 1/4, 1/4) + p_B(0, 1, 0, 0) + p_C(0, 0, 1, 0)$ for some $p_B, p_C \in [0, 1]$ with $p_0 + p_B + p_C = 1$. In DS, the belief of such reasoner is $\mu_2^{DS} = p_0(1/3, 1/3, 1/3) + (1 - p_0)(1, 0, 0)$.

First, notice that the lottery $a^{IR}_{\mu_2^{IR}} = \left(0, p_0/4; 12, p_0/4 + p_B; 13, p_0/4 + p_C; 11, p_0/4\right)$ first-order stochastically dominates the lottery $a^{DS}_{\mu_2^{DS}} = (0, p_0/3 + p_A; 12, p_0/3; 11, p_0/3)$

for all p_0, p_B and p_C. Further, the lottery $a^{IR}_{\mu_2^{IR}}$ also first-order stochastically dominates the lottery $c^{DS}_{\mu_2^{DS}} = (12, 1 - {}^{2p_0}/_3; 8, {}^{p_0}/_3; 0, {}^{p_0}/_3;)$ for all p_0, p_B and p_C. Thus, any iterative 'top-down' reasoner prefers to play IR over actions a or c in the DS game, regardless of risk preferences.[31]

Endogenous depths of reasoning. We now return to the question of whether our predictions regarding preferences over *IR* and *DS* hold even when players' types are allowed to change across games.

First, recall that any Lk type with $k \geq 2$ strictly prefers *IR* to *DS*, as they value *IR* strictly more than 12, while all Lk types value *DS* weakly less than 12. Therefore, any changes in the reasoning level that remain consistent with $k \geq 2$ in *IR* still support our main prediction.

A potential caveat arises if a player is an Lk type with $k \geq 2$ in *DS* but shifts to a type with $k \leq 1$ in *IR*. In such a case, it becomes possible for the player to value *IR* less than *DS*. However, this scenario appears unlikely for two reasons.

First, there is no evidence suggesting that players move from type Lk with $k \geq 1$ to $k = 0$ (Georganas, Healy, and Weber, 2015; Alaoui and Penta, 2016), so such a change would constitute an extreme form of holding an incorrect belief. Second, such a change in the level of reasoning is not predicted by the endogenous depth of reasoning model. In particular, when this model is applied as in Alaoui and Penta (2016; 2022) and Alaoui, Janezic, and Penta (2022, 2025), it predicts that if a player is an Lk type with $k \geq 2$ in *DS*, they will have a weakly higher level of reasoning in *IR*.[32]

[31] The only potential caveat here is that there may be an iterative 'top-down' reasoner who is extremely risk seeking *and* at the same time very pessimistic about the rationality of others (high p_0), and as such prefers the lottery $b^{DS}_{\mu_2^{DS}} = (5, {}^{p_0}/_3; 13, 1 - {}^{2p_0}/_3; 0, {}^{p_0}/_3)$ over any lotteries induced by IR. Such choices are extremely rare in our data. Of 470 choices in total, only 8 participants choose to play b in DS and value the game at $13 \leq v \leq 13.25$. As in the analysis presented in Section 4, if we control for such players by focusing on those who play c in DS, the iterative 'top-down' model of reasoning makes the unambiguous prediction that such players rank IR above DS.

[32] Following Alaoui and Penta (2016; 2022) and Alaoui, Janezic, and Penta (2022, 2025) we assume that a player's value of reasoning is determined by the maximal gain value function and that their costs of reasoning are fixed across games.

6 Concluding Remarks

In iterative reasoning models, each player best-responds to the belief that other players reason to some finite level. In this paper, we propose a novel behavior that captures the idea that players may believe that others are rational, yet cannot model their behavior. Within the prism of the level-k model, it encompasses a situation where a player believes that their opponent can reason to a higher level than they do. We present a novel experimental design that permits us to identify such behavior, and evaluate it experimentally.

We find that approximately half of the participants made choices inconsistent with a very general and permissive model of iterative 'top-down' reasoning. This is true especially if they believe that their opponents are rational. Among those, approximately two-thirds behave inconsistently with the iterative 'top-down' model.

To conclude, we provide experimental evidence that behavior may fall outside an iterative 'top-down' model of reasoning, yet players may still use alternative models, which rely on belief in their opponent's rationality, to reason and choose optimal strategies in games. Our findings support an epistemic approach that relies on explicitly modeling strategic uncertainty, to supplement the existing approach that attempts to capture it implicitly through modifications of the iterative 'top-down' models.

References

AGRANOV, M., E. POTAMITES, A. SCHOTTER, AND C. TERGIMAN (2012): "Beliefs and Endogenous Cognitive Levels: An Experimental Study," *Games and Economic Behavior*, 75(2), 449–463.

ALAOUI, L., K. JANEZIC, AND A. PENTA (2020): "Reasoning About Others' Reasoning," *Journal of Economic Theory*, 189, 105091.

——— (2025): "Coordination and Sophistication," *Journal of the European Economic Association, forthcoming.*

ALAOUI, L., AND A. PENTA (2016): "Endogenous Depth of Reasoning," *The Review of Economic Studies*, 83(4), 1297–1333.

——— (2022): "Cost-Benefit Analysis in Reasoning," *Journal of Political Economy*, 130(4), 881–925.

ARAD, A. (2012): "The Tennis Coach Problem: A Game-Theoretic and Experimental Study," *The BE Journal of Theoretical Economics*, 12(1).

ARAD, A., AND S. PENCZYNSKI (2024): "Multi-Dimensional Reasoning in Competitive Resource Allocation Games: Evidence from Intra-Team Communication," *Games and Economic Behavior*, 144(2), 355–377.

ARAD, A., AND A. RUBINSTEIN (2012a): "The 11-20 Money Request Game: A level-k Reasoning Study," *The American Economic Review*, 102(7), 3561–3573.

——— (2012b): "Multi-Dimensional Iterative Reasoning in Action: The Case of the Colonel Blotto Game," *Journal of Economic Behavior and Organization*, 84(2), 571–585.

BATTIGALLI, P., S. CERREIA-VIOGLIO, F. MACCHERONI, AND M. MARINACCI (2016): "A Note on Comparative Ambiguity Aversion and Justifiability," *Econometrica*, 84(5), 1903–1906.

Battigalli, P., and M. Siniscalchi (2002): "Strong Belief and Forward Induction Reasoning," *Journal of Economic Theory*, 106(2), 356–391.

Bernheim, D. (1984): "Rationalizable Strategic Behavior," *Econometrica*, 52(4), 1007–1028.

Brandenburger, A., and E. Dekel (1987): "Rationalizability and Correlated Equilibria," *Econometrica*, 55(6), 1391–1402.

Brandenburger, A., A. Friedenberg, and T. Kneeland (2020): "Two Approaches to Iterated Reasoning in Games," *Working Paper*.

Camerer, C., T.-H. Ho, and J.-K. Chong (2004): "A Cognitive Hierarchy Model of Games," *The Quarterly Journal of Economics*, 119(3), 861–898.

Chen, D., M. Schonger, and C. Wickens (2016): "oTree-An Open-Source Platform for Laboratory, Online, and Field Experiments," *Journal of Behavioral and Experimental Finance*, 9, 88–97.

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

Costa-Gomes, M., and V. Crawford (2006): "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," *The American Economic Review*, 96(5), 1737–1768.

Costa-Gomes, M., V. Crawford, and B. Broseta (2001): "Cognition and Behavior in Normal-Form Games: An Experimental Study," *Econometrica*, 69(5), 1193–1235.

Crawford, V. (2008): *Look-Ups as the Windows of the Strategic Soul: Studying Cognition via Information Search in Game Experiments*. In A. Caplin, and A. Schotter (Eds.), Perspectives in the Future of Economics: Positive and Normative Foundations, 249–280. Oxford.

Crawford, V., M. Costa-Gomes, and N. Iriberri (2013): "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications," *Journal of Economic Literature*, 51(1), 5–62.

Dekel, E., and A. Wolinsky (2003): "Rationalizable Outcomes of Large Private-Value First-Price Discrete Auctions," *Games and Economic Behavior*, 43(2), 175–188.

Friedenberg, A. (2019): "Bargaining Under Strategic Uncertainty: The Role of Second-Order Optimism," *Econometrica*, 87(6), 1835–1865.

Georganas, S., P. J. Healy, and R. A. Weber (2015): "On the Persistence of Strategic Sophistication," *Journal of Economic Theory*, 159, 369–400.

Gill, D., and V. Prowse (2016): "Cognitive Ability, Character Skills, and Learning to Play Equilibrium: A Level-k Analysis," *Journal of Political Economy*, 124(6), 1619–1676.

Greiner, B. (2015): "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE," *Journal of the Economic Science Association*, 1(1), 114–125.

Heinemann, F., R. Nagel, and P. Ockenfels (2009): "Measuring Strategic Uncertainty in Coordination Games," *The Review of Economic Studies*, 76, 181–221.

Kneeland, T. (2015): "Identifying Higher-Order Rationality," *Econometrica*, 83(5), 2065–2079.

Kosenkova, L. (2019): "Non-Parametric Inference in Asymmetric First-Price Auctions with k-Rationalizable Beliefs," *Working Paper*.

Nagel, R. (1995): "Unraveling in Guessing Games: An Experimental Study," *The American Economic Review*, 85(5), 1313–1326.

Nagel, R., A. Brovelli, F. Heinemann, and G. Coricelli (2018): "Neural Mechanisms Mediating Degrees of Strategic Uncertainty," *Social Cognitive and Affective Neuroscience*, 13(1), 52–62.

Stahl, D., and P. Wilson (1994): "Experimental Evidence on Players' Models of Other Players," *Journal of Economic Behavior and Organization*, 25(3), 309–327.

——— (1995): "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 10(1), 218–254.

Tan, T., and S. da Costa Werlang (1988): "The Bayesian Foundations of Solution Concepts of Games," *Journal of Economic Theory*, 45(2), 370–391.

Weinstein, J. (2016): "The Effect of Changes in Risk Attitude on Strategic Behavior," *Econometrica*, 84(5), 1881–1902.

Wright, J., and K. Leyton-Brown (2014): "Level-0 Meta-Models for Predicting Human Behavior in Games," *ACM Conference on Economics and Computation (ACM-EC).*