

Schipper, Burkhard; Zhou, Hang

Working Paper

Extensive-form level-k thinking

Working Paper, No. 352

Provided in Cooperation with:

University of California Davis, Department of Economics

Suggested Citation: Schipper, Burkhard; Zhou, Hang (2022) : Extensive-form level-k thinking, Working Paper, No. 352, University of California, Department of Economics, Davis, CA

This Version is available at:

<https://hdl.handle.net/10419/267016>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

EXTENSIVE-FORM LEVEL- k THINKING*

Burkhard C. Schipper[†] Hang Zhou[‡]

September 1, 2022

Abstract

Level- k thinking and Cognitive Hierarchy have been widely applied as a normal-form solution concept in behavioral and experimental game theory. We consider the extension of level- k thinking to extensive-form games. Player's may learn about levels of opponents' thinking during the play of the game because some information sets may be inconsistent with certain levels. In particular, for any information set reached, a level- k player attaches the maximum level- ℓ thinking for $\ell < k$ to her opponents consistent with the information set. We compare extensive-form level- k thinking with other solution concepts such as level- k thinking in the associated normal-form, extensive-form rationalizability, Δ -rationalizability, iterated admissibility, and backward induction. We use extensive-form level- k thinking to reanalyze data from some prior experiments in the literature.

Keywords: Level- k thinking, Cognitive hierarchy, Theory-of-Mind, Rationalizability, Iterated admissibility, Extensive-form rationalizability, Δ -rationalizability, Mutual belief in rationality, Experimental game theory.

JEL-Classification: C72, C92, D91.

*We thank Dieter Balkenborg, Russell Cooper, Douglas De Jong, Piotr Evdokimov, Rosemarie Nagel, Thomas Ross and Aldo Rustichini for sharing their data with us. We also thank participants in the Bay Area Experimental Workshop, 2019, for helpful comments.

[†]Department of Economics, University of California, Davis. Email: bcschipper@ucdavis.edu

[‡]School of Finance, Shanghai University of Finance and Economics. Email: zhouhang@mail.shufe.edu.cn

1 Introduction

The core idea of level- k thinking or iterated reasoning about other players is as old as game theory and at the heart of strategic reasoning. Even before the seminal “Theory of Games and Economic Behavior” had been published by von Neumann and Morgenstern in 1944, Morgenstern (1928, p. 98) emphasized it in his work on predictions in social situations:¹

“Sherlock Holmes, pursued by his opponent, Moriarty, leaves for Dover. The train stops at a station on the way, and he alights there rather than traveling on to Dover. He has seen Moriarty at the railway station, recognizes that he is very clever, and expects that Moriarty will take a special faster train in order to catch him at Dover. Holmes’ anticipation turns out to be correct. But what if Moriarty had been still more clever, had estimated Holmes’ mental abilities better and had foreseen his actions accordingly? Then obviously he would have traveled to the intermediate station. Holmes, again, would have had to calculate that, and he himself would have decided to go on to Dover. Whereupon Moriarty would have “reacted” differently. Because of so much thinking, they might not have been able to act at all or the intellectually weaker of the two would have surrendered to the other in the Victoria Station, since the whole flight would have become unnecessary. Examples of this kind can be drawn from everywhere.”

It has been studied in various different forms as sequential best response learning (Cournot, 1838), hierarchies of beliefs (Harsanyi, 1967), iterated admissibility and iterated dominance (Farquharson, 1969, Brams, 1975, Moulin, 1979), rationalizability (Spohn, 1982, Bernheim, 1984, Pearce, 1984), k -level mutual belief in rationality and variants thereof (Tan and Werlang, 1988, Battigalli and Siniscalchi, 2002, Brandenburger, Friedenberg, Keisler, 2008, Heifetz, Meier, and Schipper, 2019), level- k thinking (Nagel, 1995), and cognitive hierarchies (Stahl and Wilson, 1994, 1995, Camerer, Ho, Chong, 1994). Latter work on level- k thinking was very much inspired by experiments and has been applied to a wide variety of experimental games (see Crawford, Costa-Gomez, Iriberry, 2013, for a survey) and seen applications to auctions (Crawford and Iriberry, 2007), mechanism design (Kneeland, 2022), financial market microstructure (Zhou, 2022),

¹While the origin of level- k thinking is often traced back to the Beauty Contest described by Keynes in 1936, we note that Morgenstern’s work predates Keynes.

and even general equilibrium (Carvajal and Zhou, 2022). We learn from the literature that details matter. In this paper, we focus on details that arise in extensive-form games. Moreover, we focus on level- k thinking as currently this is a dominant solution concept in experimental game theory thought to capture some notion of bounded rationality.

Although level- k reasoning has been applied to extensive-form games (e.g., Kawagoe and Takizawa, 2012, Ho and Su, 2013, Garcia-Pola, Iriberry, and Kovarik, 2020), it is a solution concept for normal-form games. Players hold a first-level belief over opponent's behavior (often called level-0 types). A level-1 player best responds to such a first-level belief. A level-2 player best response to his belief that others are level-1 players. A level-3 player best responds to level-2 players etc. The issue is that in extensive-form games, players may learn about the opponents' levels of thinking during the play because some information sets cannot be reached when opponents use certain levels of thinking. Any level of opponents' reasoning that a player learns during play must be below her own level of reasoning. Information about the opponents' level of reasoning is useful for predicting opponents' future play.

In this paper, we present an extensive-form version of level- k thinking that allows for updating of beliefs over opponents' levels of thinking in extensive-form games. We compare extensive-form level- k thinking with normal-form level- k thinking, extensive-form rationalizability, strong Δ -rationalizability, iterated admissibility, and backward induction. We focus on the comparison with these solution concepts because all of them can be interpreted as some form of iterative reasoning. We show that for initial full-support beliefs, extensive-form level- k thinking refines normal-form level- k thinking outcomes. However, while normal-form level- k thinking refines level- k normal-form rationalizability, extensive-form level- k thinking does not refine (and is not refined by) k -level extensive-form rationalizability. We also show that extensive-form level- k thinking differs from k -level Δ -rationalizability, k -level iterative admissibility, and k -level backward induction.

Our goal in proposing a notion of extensive-form level- k thinking is not to put up another contender in a horse race of solution concepts that magically predicts behavioral data better than any other in games in extensive form. Rather, our hope is that by confronting experimental data with various solution concepts such as extensive-form level- k thinking, normal-form level- k thinking, and k -level extensive-form rationalizability, we are able to learn about particular features of human strategic reasoning that are reflected in one solution concept but not in others keeping other features fixed. For instance, by comparing the fit of extensive-form level- k thinking and normal-form level- k thinking, we can learn about the prevalence of forward induction *given* comparable levels of reasoning

and assumptions on level-1 beliefs/level-0 behavior. Or by comparing the fit of extensive-form level- k thinking and k -level extensive-form rationalizability, we can learn about the impact of assumptions on level-1 beliefs/level-0 behavior *given* comparable levels of reasoning and the ability to do forward induction. As a first proof of concept, we reanalyze data from versions of the battle-of-the-sexes game with an outside option by Cooper et al. (1993), Balkenborg and Nagel (2016), and Evdokimov and Rustichini (2016).

We are not aware of any other paper that presents an extensive-form version of level- k thinking. The closest paper to ours is Ho and Su (2013) who consider repeated play of an extensive-form game. Players use level- k thinking applied to each stage game of the repeated play. Yet, they allow for updating of levels only between stages of the repeated play. Although Ho and Su (2013) study repeated extensive-form games, they apply normal-form level- k thinking to the normal-form of the stage game with updating *between* repeated play of the stage game. This is different from our approach since we allow for updating about opponents' levels of thinking *within* the play of the extensive-form game. That is, they focus on learning between stage games while we allow for dynamic level- k thinking even within one-shot play of extensive-form games. This makes our solution concept applicable even to one-shot play of games in extensive form. Feng and Wang (2019) also study learning of levels of reasoning for both, normal-form level- k thinking and cognitive hierarchy in repeated play of normal-form games.

In our notion of extensive-form level- k thinking, players are able to learn about opponents' levels of reasoning from opponents' play throughout the game. Consequently, their own play may vary with what they learned about opponents' levels of reasoning earlier in the game. There is quite some experimental evidence for the assumption that a player's behavior depend on her belief about the levels of reasoning of opponents; see for instance Agranov et al. (2012) and Alaoui, Janezic, and Penta (2020). The assumption is also consistent with the idea that the levels of reasoning displayed by a player might be endogenous (Alaoui and Penta, 2016).

The paper is organized as follows: The next section recalls level- k thinking in the normal-form and compares it to rationalizability. This sets the stage for Section 3 that introduces the definition of extensive-form level- k thinking, compares it to level- k thinking in the normal-form, extensive-form rationalizability, strong Δ -rationalizability, iterated admissibility, and backward induction. In Section 4, we present a simply reanalysis of data from prior experiments on games. Proofs are relegated to the appendix.

2 Normal-Form Level-k Thinking

First, we review level- k thinking for games in normal-form. This will turn out to be useful when comparing it to extensive-form level- k thinking. We consider finite games in normal-form $\langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$ that consist of a nonempty finite set of players $N = \{1, \dots, n\}$ and for each player $i \in N$, a nonempty finite set of actions A_i and a utility function $u_i : A \rightarrow \mathbb{R}$ with $A := \times_{i \in N} A_i$. As usual, for any player $i \in N$, we denote by $A_{-i} := \times_{j \in N \setminus \{i\}} A_j$ the set of action profiles of player i 's opponents.² Denote by $\Delta(A_{-i})$ the set of probability measures on A_{-i} . A belief of player i over opponents' actions is denoted by $\beta_i \in \Delta(A_{-i})$.

We say that player i 's action $a_i \in A_i$ is *rational* with β_i if a_i maximizes player i 's expected utility with β_i . With these definitions in place, we can define the by now standard solution concept of level- k thinking³ that has been widely applied in experimental game theory.

Definition 1 (Normal-Form Level- k Thinking) Fix a first level belief profile $\beta^1 = (\beta_i^1)_{i \in N}$ with $\beta_i^1 \in \Delta(A_{-i})$ for each $i \in N$. Define inductively for each player $i \in N$,

$$\begin{aligned} B_i^1(\beta^1) &= \{\beta_i^1\} \\ L_i^1(\beta^1) &= \{a_i \in A_i : a_i \text{ is rational for player } i \text{ with belief } \beta_i^1\} \\ &\vdots \\ B_i^k(\beta^1) &= \{\beta_i \in \Delta(A_{-i}) : \beta_i(L_{-i}^{k-1}(\beta^1)) = 1\} \\ L_i^k(\beta^1) &= \left\{ a_i \in A_i : \begin{array}{l} \text{There exists } \beta_i \in B_i^k(\beta^1) \text{ such that } a_i \\ \text{is rational for player } i \text{ with belief } \beta_i. \end{array} \right\}. \end{aligned}$$

For $k \geq 1$, we call $L_i^k(\beta^1)$ the set of (normal-form) level- k thinking actions of player i anchored by the profile of first-level beliefs β^1 .

Several remarks are in order: First, the first-level belief β^1 is often interpreted as behavior of level-0 players. Since it is also often assumed that there are no actual level-0 players, it is appropriate to interpret it as a (fictitious) belief over opponents' play.

Second, level- k thinking is not one solution concept but a collection of solution concepts, one for each first-level belief β^1 / level-0 behavior and each level k . In applications,

²For games in normal form, we will make use of the “ $-i$ ” notational convention for any objects indexed by players.

³Although we use the established terminology, we must admit that we do not know how level- k thinking is related to actual thought processes in the human brain.

the first-level belief is often fixed to uniform distribution (e.g., Nagel, 1995) although in some applications other distributions are more natural (e.g., Arad and Rubinstein, 2012). These assumptions seem to reflect assumptions of either unpredictable behavior (i.e., in the spirit of the principle of insufficient reason) or non-strategic level-0 behavior.

Third, when iterated best responses are unique, i.e., the sets $L_i^k(\beta^1)$, $k \geq 1$, are singleton, then also $B_i^{k+1}(\beta^1)$, $k \geq 1$, are singleton. In such a case, it does not matter whether or not we allow for correlated beliefs or, alternatively, would restrict to independent beliefs. In applications, unique best responses are frequently assumed in the literature.

2.1 Level-k Thinking versus Rationalizability

The idea of using iterated reasoning about opponents play in a solution concept featured already in solution concepts developed earlier in game theory. In particular, rationalizability introduced by Spohn (1982), Bernheim (1984), and Pearce (1984) is defined inductively using player's beliefs about rational actions of opponents.⁴

Definition 2 (Rationalizability) *Define inductively for each player $i \in N$,*

$$\begin{aligned}
 B_i^1 &= \Delta(A_{-i}) \\
 R_i^1 &= \left\{ a_i \in A_i : \begin{array}{l} \text{There exists } \beta_i \in B_i^1 \text{ such that } a_i \\ \text{is rational for player } i \text{ with belief } \beta_i. \end{array} \right\} \\
 &\vdots \\
 B_i^k &= \{ \beta_i \in \Delta(A_{-i}) : \beta_i(R_{-i}^{k-1}) = 1 \} \\
 R_i^k &= \left\{ a_i \in A_i : \begin{array}{l} \text{There exists } \beta_i \in B_i^k \text{ such that } a_i \\ \text{is rational for player } i \text{ with belief } \beta_i. \end{array} \right\}
 \end{aligned}$$

For any $k \geq 1$, we call R_i^k the set of k -level rationalizable actions of player i . The set of player i 's rationalizable actions is

$$R_i^\infty = \bigcap_{k=1}^{\infty} R_i^k$$

It is well-known that rationalizability is strategy-equivalent to iterated elimination of strictly dominated actions (see Pearce, 1984). In fact, for any level $k \geq 1$, level- k rationalizability (i.e., rationalizability up to level k) is strategy-equivalent to k -iterated

⁴Like most of the literature, we focus on correlated rationalizability; see Brandenburger and Dekel, 1987, and Tan and Werlang, 1988, for more on rationalizability.

elimination of strictly dominated actions. An action is strictly dominated if there exists a possibly mixed action that yields a strictly higher expected utility no matter what opponents play. Thus, rationalizability does not only provide a prediction in the limit when k goes to infinity, but also for every finite level k . This has been previously used in experiments to partially identify levels of beliefs (see for a similar approach, Li and Schipper, 2020). In our context, it is now natural to ask about how behavior implied by level- k thinking is related to level- k rationalizable actions in games in normal form. For any first-order belief (i.e., any level-0 behavior), the level- k behavior is k -level rationalizable. That is, level- k thinking implies k -level rationalizability.

Proposition 1 *For any finite game in normal-form, $\langle N, (A_i)_{i \in I}, (u_i)_{i \in I} \rangle$, any profile of first-level beliefs, β^1 , and level $k \geq 1$, $L^k(\beta^1) \subseteq R^k$.*

While obvious at the first level, the proof by induction in the appendix reveals that at any level k , the belief that a player following level k thinking entertains about opponents' play is a belief that k -level rationalizes her action. That is, the inclusion holds not only for actions but also beliefs.

The most obvious difference between rationalizability and level- k thinking is that rationalizability does not fix first-level beliefs. Initially, it allows for *any* first-level beliefs over opponents' actions. This is useful when there is no "natural" first-level belief and when it is reasonable to assume that players could entertain any initial belief. Based on Proposition 1, one may be tempted to claim that level- k reasoning yields sharper predictions than rationalizability. Yet, level- k thinking does not explain the first-level beliefs. That is, it does not provide a theory of first-level beliefs or level-0 behavior. Rather, *given* a first-level belief of players / assumption of level-0 behavior of the analyst that, while extremely useful, is necessarily ad hoc because it is outside the model of level- k thinking, it yields a sharper prediction than if the analyst considers any first-level belief as in level- k rationalizability.

In order to emphasize that (1) first-level beliefs of level- k thinking yield strong refinement power, and (2), that level- k thinking differs from rationalizability just in the first-level belief, we state the following weaker converse to Proposition 1. For every rationalizable action there exists a first-level belief with which the action is rationalizable with level-1 thinking. Note that this implies that we can find first-level beliefs that justify any arbitrarily high k -level rationalizable action with level-1 thinking. Thus, ex ante tying the hands to a particular first-level belief/assumption of level-0 behavior is crucial in

experiments if level- k thinking is to have predictive power beyond rationalizability and the identification of levels is to be meaningful.

Proposition 2 *For any finite game in normal-form, $\langle N, (A_i)_{i \in I}, (u_i)_{i \in I} \rangle$, every player $i \in N$, and every $a_i \in R_i^\infty$, there exists a first-level belief $\beta^1 \in \Delta(A_{-i})$ such that $a_i \in L_i^1(\beta^1)$.*

While the observation might be obvious to some, we present a short proof in the appendix.

It is also worth emphasizing another difference between level- k thinking and rationalizability: Latter is a reduction procedure on beliefs that implies a reduction procedure on actions while for former this is not necessarily the case. This demonstrated in the following simple example:

Example 1 *Consider the following the following variant of the matching pennies game:*

$$\begin{array}{c} \\ U \end{array} \begin{array}{cc} L & R \\ \left(\begin{array}{cc} 2, -1 & -1, 2 \\ -1, 1 & 1, -1 \end{array} \right) \end{array}$$

Apply level- k thinking anchored with a uniform first-level belief for each player. Then the sets of level- k thinking actions are derived subsequently by

	<i>Row player</i>	<i>Column player</i>
$L_i^1(\beta^1)$	$\{U\}$	$\{R\}$
$L_i^2(\beta^1)$	$\{D\}$	$\{R\}$
$L_i^3(\beta^1)$	$\{D\}$	$\{L\}$
$L_i^4(\beta^1)$	$\{U\}$	$\{L\}$
$L_i^4(\beta^1)$	$\{U\}$	$\{R\}$
\vdots	\vdots	\vdots

We note that level- k thinking results in a choice cycle. So clearly, it is not the case that the set of level- k actions are refined level by level.

The fact that level- k thinking is not necessarily a reduction procedure is certainly not desirable from an epistemic point of view. Yet, it matches curiously with the reasoning reflected in the quote from Morgenstern (1928) that we stated in the introduction.

3 Extensive-Form Level-k Thinking

Consider a finite extensive-form game with possibly imperfect information and simultaneous moves, perfect recall but finite horizon $\langle N, H, P, (\mathcal{I}_i)_{i \in N}, (u_i)_{i \in N} \rangle$ defined by

- A nonempty finite set of players N .
- A set H of finite sequences of action profiles (i.e., histories) such that
 - $\emptyset \in H$,
 - If history $(a^m)_{m=1, \dots, M} \in H$ and $M' < M$, then also the subhistory $(a^m)_{m=1, \dots, M'} \in H$.

The set of terminal histories are histories with no successors. We denote them as usual by Z .

- A player function $P : H \setminus Z \rightarrow 2^{N \cup \{c\}} \setminus \{\emptyset\}$ that assigns to each nonterminal history $h \in H \setminus Z$ a nonempty subset of players $P(h) \subseteq N \cup \{c\}$ that may include nature, c . That is, players are allowed to move simultaneously. Moreover, nature is allowed to move any time and even simultaneously with other players.

With this notation, let $A_i(h)$ be the nonempty set of actions of player $i \in P(h)$ at the non-terminal history $h \in H \setminus Z$. Moreover, we let $a \in \times_{i \in P(h)} A_i(h)$ be the action profile of players moving at history h . That is, $h \in H \setminus Z$ and $a \in \times_{i \in P(h)} A_i(h)$, then $(h, a) \in H$.

- For each player $i \in N$, a partition \mathcal{I}_i of non-terminal histories in $H_i = \{h \in H \setminus Z : i \in P(h)\}$ at which he moves. Elements of \mathcal{I}_i are called information sets of player i . As usual, we require that for any information set $I_i \in \mathcal{I}_i$, $A_i(h) = A_i(h')$ for any $h, h' \in I_i$. Thus, we can simply denote by $A_i(I_i)$ the set of player i 's actions at information set I_i . Each player's information sets are required to satisfy perfect recall.
- For each player $i \in N$, a von Neumann-Morgenstern utility function over lotteries of terminal histories. We denote by $u_i : Z \rightarrow \mathbb{R}$ player i 's Bernoulli utility index.

See for instance Osborne and Rubinstein (1995, Chapters 6.3.2 and 11.1.2) for further details on extensive-form games including perfect recall. We allow for simultaneous moves of players. We also allow for imperfect information and moves of nature at any time during game including simultaneously with other players. We do not have to assume

a prior probability measure over moves of nature but such an assumption can be added whenever it is required.

For any player $i \in N$, a (pure) strategy of player i assigns to each of her information sets an action available at that information set. Formally, a strategy s_i is a function $s_i : \mathcal{I}_i \rightarrow \bigcup_{I_i \in \mathcal{I}_i} A_i(I_i)$ such that $s_i(I_i) \in A_i(I_i)$ for all $I_i \in \mathcal{I}_i$. Let S_i denote the set of player i 's strategies. Let $N_c := N \cup \{c\}$. We treat nature as a player with information sets that are singleton. Define $S := \times_{i \in N_c} S_i$ if nature moves in the game. Otherwise, $S := \times_{i \in N} S_i$. Similarly, for any $i \in N$, $S_{-i} := \times_{j \in N_c \setminus \{i\}} S_j$ if nature moves in the game. Otherwise, $S_{-i} := \times_{j \in N \setminus \{i\}} S_j$.

For every player $i \in N$, we say a strategy $s_i \in S_i$ reaches an information set if there exists a profile of opponents' strategies s_{-i} such that (s_i, s_{-i}) reaches this information set. Similarly, a profile of opponents' strategies s_{-i} reaches an information set if there exists a strategy s_i of player i such that (s_i, s_{-i}) reaches this information set.

A belief system of player $i \in N$

$$\bar{\beta}_i := (\bar{\beta}_i(I_i))_{I_i \in \mathcal{I}_i} \in \prod_{I_i \in \mathcal{I}_i} \Delta(S_{-i})$$

is a profile of beliefs, a belief $\bar{\beta}_i(I_i) \in \Delta(S_{-i})$ about other players' strategies for each information set $I_i \in \mathcal{I}_i$, with the following properties:

- $\bar{\beta}_i(I_i)$ reaches I_i , i.e., $\bar{\beta}_i(I_i)$ assigns probability 1 to the set of strategy profiles of the other players that reaches I_i .
- If information set I_i precedes information set I'_i , then $\bar{\beta}_i(I'_i)$ is derived from $\bar{\beta}_i(I_i)$ by conditioning whenever possible.

Denote by \bar{B}_i the set of all belief systems of player $i \in N$.

For a player i and an information set I_i , a strategy s'_i is an I_i -replacement of strategy s_i if s'_i agrees with s_i on all information sets strictly preceding I_i .

With a belief system $\bar{\beta}_i$, strategy s_i is sequentially rational for player i at the information set I_i if s_i does not reach I_i or if s_i does reach I_i but there does not exist I_i -replacement of s_i which yields a strictly higher expected utility given $\bar{\beta}_i(I_i)$ on S_{-i} .

Definition 3 (Extensive-Form Level- k Thinking) *Given a belief system of first-level*

beliefs, $\bar{\beta}^1 = (\bar{\beta}_i^1)_{i \in N}$ with $\bar{\beta}_i^1 \in B_i$, define inductively for all $i \in N$,

$$\begin{aligned}
\bar{B}_i^1(\bar{\beta}^1) &= \{\bar{\beta}_i^1\} \\
\bar{L}_i^1(\bar{\beta}^1) &= \left\{ s_i \in S_i : \begin{array}{l} \text{For every information set } I_i \in \mathcal{I}_i, s_i \text{ is sequentially} \\ \text{rational at } I_i \text{ with respect to } \bar{\beta}_i^1. \end{array} \right\} \\
&\vdots \\
\bar{B}_i^k(\bar{\beta}^1) &= \left\{ \bar{\beta}_i \in \bar{B}_i : \begin{array}{l} \text{For every information set } I_i \in \mathcal{I}_i, \text{ if there exists } \ell \\ \text{with } 1 \leq \ell < k \text{ for which there exists } s_{-i} \in \bar{L}_{-i}^\ell(\bar{\beta}^1) \\ \text{such that } s_{-i} \text{ reaches } I_i, \text{ let } \bar{\ell} \text{ be the largest such } \ell. \\ \text{Then } \bar{\beta}_i(I_i) \text{ assigns probability 1 to } \bar{L}_{-i}^{\bar{\ell}}(\bar{\beta}^1). \text{ Otherwise,} \\ \text{if there is no such } \ell, \text{ then let } \bar{\beta}_i(I_i) = \bar{\beta}_i^1(I_i). \end{array} \right\} \\
\bar{L}_i^k(\bar{\beta}^1) &= \left\{ s_i \in S_i : \begin{array}{l} \text{There exists } \bar{\beta}_i \in \bar{B}_i^k(\bar{\beta}^1) \text{ with which for every information} \\ \text{set } I_i \in \mathcal{I}_i, s_i \text{ is sequentially rational at } I_i. \end{array} \right\}
\end{aligned}$$

For any level k , we call $\bar{L}_i^k(\bar{\beta}_i)$ player i 's set of extensive-form level- k thinking strategies.

Several remarks are in order: First, a player's assumption of opponents' level-0 behavior is modelled as level-1 belief. However, we like to emphasize that since a player may move at various information sets of the game, she forms possibly different beliefs about opponents' behavior at each of her information sets.⁵ According to the notion of belief system, each such a belief must be consistent with having reached this information set and derived by conditioning whenever possible. That is, different from normal-form level- k thinking, a player not just forms beliefs about opponents' behavior before playing the game but also at each of her information sets throughout the game. The assumption of level-0 behavior of opponents is now a collection of assumptions, one belief at each of her information sets.

Second, extensive-form level- k thinking features the best rationalizability principle (Battigalli, 1996). At each of her information sets, a player following extensive-form level- k thinking assigns the highest possible level ℓ -thinking with $\ell < k$ to opponents that is consistent with reaching the information set. Intuitively, a player does not easily label opponents' behavior as a "mistake". Rather a player tries to make sense of opponents' behavior as much as it possible within her own (limited) thinking. Such an approach makes quite some sense when allowing for learning from opponents' play. If opponents' play is judged easily as a mistake, then there is not much to learn from.

⁵Kawagoe and Takizawa (2012) consider a normal-form version of level- k thinking applied to centipede games in which they allow for level-1 uniform beliefs over actions *at each* node rather than just level-1 uniform initial beliefs over strategies. This definition is already a small step away from normal-form level- k thinking towards extensive-form level- k thinking.

Third, a player’s belief is updated to lower levels of opponents’ thinking along the path of play whenever information sets are reached that cannot be reached with strategies of opponents’ featuring higher levels of thinking. Of course, this depends on the player’s own level of thinking since one defining feature of level- k thinking is a form of naiveté, namely that a player with level k must believe opponents’ feature levels of thinking strictly lower than level k .

Fourth, the phrasing of the definition of $\bar{B}_i^k(\bar{\beta}^1)$ appears somewhat awkward at the first glance, since it involves an “if ... then ... Otherwise ...” clause. Partially, this captures the best rationalizability principle above (i.e., the “then” statement). Yet, the “if” and “Otherwise” clauses emphasize that an information set may not be reachable with any extensive-form level- ℓ thinking strategy, for $\ell < k$. What shall a player believe in such a case? We assume that in such a case, the player resorts nevertheless to his first-level belief. Such an assumption is innocuous if the first-level belief is a full support belief like for instance uniform belief as often assumed in the literature on level- k thinking.

In the following subsections, we explore extensive-form level- k thinking by comparing it to various other solution concepts.

3.1 Normal-Form versus Extensive-Form Level- k Thinking

The key difference between level- k thinking in the normal form and extensive-form level- k thinking is that players can update beliefs about opponent’s level- ℓ thinking, for $\ell < k$, conditional on information sets reached. At each information set reached, player i with level- k thinking attributes the highest level- ℓ thinking, $\ell < k$, to opponents consistent with the information set. It embodies the best rationalization principle (Battigalli, 1996). To what extent does the best rationalizability principle matter? First, we show that, if $\bar{\beta}^1$ is a profile of full-support belief systems, then extensive-form level-1 strategies are equivalent to normal-form level-1 strategies.

Proposition 3 *Consider any finite extensive-form game with perfect recall. Let $\bar{\beta}^1$ be a profile of full-support belief systems and β^1 a profile of full-support beliefs in the associated normal-form consistent with $\bar{\beta}^1$. Then for any player $i \in N$, $\bar{L}_i^1(\bar{\beta}^1) = L_i^1(\beta^1)$.*

The observation implies in particular that if beliefs are uniform both in the extensive-form game and the associated normal-form (as frequently assumed in experimental work), then level-1 thinking strategies coincide.

Table 1: Solutions to various games

Reny (1992) game	Extensive-form level- k (uniform level-1 belief system)		Normal-form level- k (uniform level-1 belief)		k -level extensive-form rationalizability		Backward induction	
	Player 1	Player 2	Player 1	Player 2	Player 1	Player 2	Player 1	Player 2
1	$\{(O_1, *)\}$	$\{(C_2, o_2)\}$	$\{(O_1, *)\}$	$\{(C_2, o_2)\}$	$\{(O_1, *), (C_1, c_1)\}$	$\{(O_2, *), (C_2, o_2)\}$	S_1	$\{(O_2, o_2), (C_2, o_2)\}$
2	$\{(O_1, *)\}$	$\{(C_2, o_2)\}$	$\{(O_1, *)\}$	S_2	$\{(O_1, *), (C_1, c_1)\}$	$\{(C_2, o_2)\}$	$\{(O_1, o_1), (C_1, o_1)\}$	$\{(O_2, o_2), (C_2, o_2)\}$
3	$\{(O_1, *)\}$	$\{(C_2, o_2)\}$	$\{(O_1, *), (C_1, c_1)\}$	S_2	$\{(O_1, *), (C_1, c_1)\}$	$\{(C_2, o_2)\}$	$\{(O_1, o_1), (C_1, o_1)\}$	$\{(O_2, o_2)\}$
4	$\{(O_1, *)\}$	$\{(C_2, o_2)\}$	$\{(O_1, *), (C_1, c_1)\}$	S_2	$\{(O_1, *), (C_1, c_1)\}$	$\{(C_2, o_2)\}$	$\{(O_1, o_1)\}$	$\{(O_2, o_2)\}$
5	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots

HMS game	Extensive-form level- k (uniform level-1 belief system)		Normal-form level- k (uniform level-1 belief)		k -level extensive-form rationalizability		Backward induction	
	Player 1	Player 2	Player 1	Player 2	Player 1	Player 2	Player 1	Player 2
1	$\{(O_1, *), (C_1, c_1)\}$	$\{(O_2, *)\}$	$\{(O_1, *), (C_1, c_1)\}$	$\{(O_2, *)\}$	$\{(O_1, *), (C_1, c_1)\}$	S_2	S_1	S_2
2	$\{(O_1, *), (C_1, c_1)\}$	$\{(C_2, *)\}$	S_1	$\{(C_2, *)\}$	$\{(O_1, *), (C_1, c_1)\}$	$\{(C_2, *)\}$	S_1	S_2
3	$\{(O_1, *), (C_1, c_1)\}$	$\{(C_2, *)\}$	$\{(O_1, *), (C_1, c_1)\}$	S_2	$\{(O_1, *), (C_1, c_1)\}$	$\{(C_2, *)\}$	S_1	S_2
4	$\{(O_1, *), (C_1, c_1)\}$	$\{(C_2, *)\}$	$\{(O_1, *), (C_1, c_1)\}$	$\{(C_2, *)\}$	$\{(O_1, *), (C_1, c_1)\}$	$\{(C_2, *)\}$	S_1	S_2
5	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots

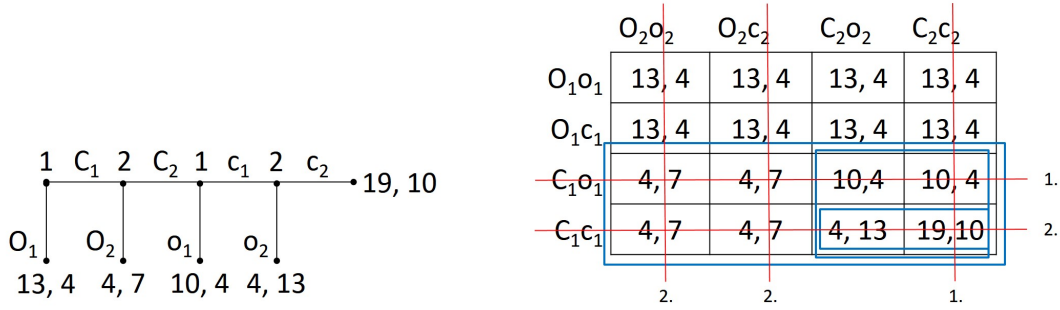
Centipede game	Extensive-form level- k (uniform level-1 belief system)		Normal-form level- k (uniform level-1 belief)		k -level extensive-form rationalizability		Backward induction	
	Player 1	Player 2	Player 1	Player 2	Player 1	Player 2	Player 1	Player 2
1	$\{(C_1, c_1)\}$	$\{(C_2, o_2)\}$	$\{(C_1, c_1)\}$	$\{(C_2, o_2)\}$	S_1	$\{(O_2, *), (C_2, o_2)\}$	S_1	$\{(O_2, o_2), (C_2, o_2)\}$
2	$\{(O_1, o_1)\}$	$\{(C_2, o_2)\}$	$\{(C_1, o_1)\}$	$\{(C_2, o_2)\}$	$\{(O_1, *), (C_1, o_1)\}$	$\{(O_2, *), (C_2, o_2)\}$	$\{(O_1, o_1), (C_1, o_1)\}$	$\{(O_2, o_2), (C_2, o_2)\}$
3	$\{(C_1, o_1)\}$	$\{(O_2, *)\}$	$\{(C_1, o_1)\}$	$\{(O_2, *)\}$	$\{(O_1, *), (C_1, o_1)\}$	$\{(O_2, *)\}$	$\{(O_1, o_1), (C_1, o_1)\}$	$\{(O_2, o_2)\}$
4	$\{(O_1, *)\}$	$\{(O_2, *)\}$	$\{(O_1, *)\}$	$\{(O_2, *)\}$	$\{(O_1, *)\}$	\vdots	$\{(O_1, o_1)\}$	$\{(O_2, o_2)\}$
5	\vdots	\vdots	\vdots	S_2	\vdots	\vdots	\vdots	\vdots
6	\vdots	\vdots	S_1	S_2	\vdots	\vdots	\vdots	\vdots

* refers to any action at the player's second information set.

The observation does not extend to level-2 thinking strategies as the following example demonstrates:

Example 2 (Extensive-form level- k strategies refine normal-form level- k strategies for $k \geq 2$) Consider a version of a game in Figure 1 due to Reny (1992) (which is itself a variant of a centipede game). To the left, we print the game in extensive form; to the right the associated normal-form.⁶ The extensive-form level- k thinking strategies and

Figure 1: Reny (1992) Game



the normal-form level- k thinking strategies are printed in Table 1 for any $k \geq 1$. (We let “*” stand for any action.) We observe that at any level $k \geq 2$, the extensive-form level- k strategies of player 2 are a subset of the normal-form level- k strategies. For $k \geq 3$, this holds not only for strategies of player 2 but even outcomes. The reason is that once player 2 gets to move (i.e., reaches her first information set), she is certain at level 2 that player 1 does not use extensive-form level-1 thinking since any extensive-form level-1 thinking strategy of player 1 prescribes O_1 at the root of the tree. She must now think that player 1 behaves uniformly over her strategies and with such a belief her strategy (C_2, o_2) is uniquely rational. In contrast, normal-form level-2 thinking of player 2 presumes level-1 thinking strategies of player 1, i.e., any strategy in $\{(O_1, o_1), (O_1, c_1)\}$. Thus, player 2 is indifferent among all her strategies and hence any strategy in S_2 is consistent with normal-form level-2 thinking.

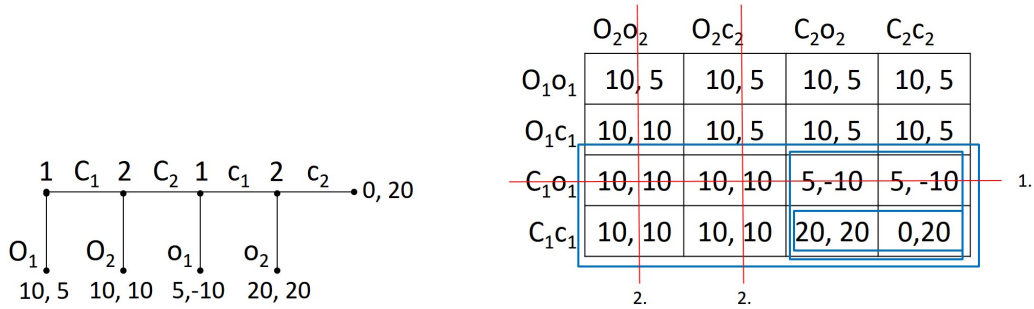
Above example shows that at level 2, strategies only differ between extensive-form level- k thinking and normal-form level- k thinking but not outcomes. One could argue that this fact is behaviorally irrelevant at level 2 as only outcomes “should” matter. Note

⁶Ignore for the time being the red lines and blue boxes in the normal form of Figure 1. They pertain to extensive-form rationalizability and will be discussed later.

though that strategies could be elicited via the strategy method (Selten, 1967) so that not just outcomes of the extensive-form game are behaviorally relevant. Nevertheless, even if one takes the view that only outcomes matter, it should be noted that in this example outcomes differ at level $k \geq 3$ as well. Moreover, the next example shows that already at level 2, outcomes of extensive-form level- k thinking can differ from normal-form level- k thinking.

Example 3 (Extensive-form level-2 outcomes refine normal-form level-2 outcomes) Consider the game in Figure 2 from Heifetz, Meier, and Schipper (2021) (i.e., HMS game for short): The extensive-form level- k thinking strategies and the normal-form

Figure 2: HMS Game



level- k thinking strategies (both with uniform level-1 beliefs) are printed in Table 1 for any $k \geq 1$. We observe that at level 2, the extensive-form level-2 thinking strategies are $\{(O_1, *), (C_1, c_1)\}$ and $\{(C_2, *)\}$ for players 1 and 2, respectively, while for normal-form level-2 thinking they are S_1 and $\{(C_2, *)\}$ for players 1 and 2, respectively. (Again, we let “*” stand for any action.) In particular, normal-form level-2 thinking allows outcome $(5, -10)$ to emerge while this is ruled out with extensive-form level-2 thinking. The reason is that under normal-form level-2 thinking, player 2 is indifferent among all strategies in the normal-form since he believes that player 2 plays O_2 . In contrast, in the extensive-form, when player 1 reaches his second information set, he is certain that player 2 does not follow extensive-form level-1 thinking but must be “level-0”. Consequently, at his second information set, he has uniform beliefs about the actions of player 2 at the last information set, with which only c_1 is rational.

Both examples beg the question whether or not in general for any level $k \geq 1$, extensive-form level- k thinking is a (weak) refinement (both in terms of strategies and

outcomes) of normal-form level- k thinking. The examples suggest this to be the case. More generally, we can show that for any full-support initial belief β^1 and level k , the set of outcomes reached by extensive-form level- k thinking refines the set of outcomes reached by normal-form level- k thinking.

To state this assertion more formally, we require the following definition. For any strategy profile $s \in S$, let $z(s)$ denote the terminal history reached by s . For any nonempty subset of strategy profiles $S' \subseteq S$, let $Z(S') = \{z \in Z : z = z(s), s \in S'\}$. Note that for any nonempty $S', S'' \in S$, $S' \subseteq S''$ implies $Z(S') \subseteq Z(S'')$.

Proposition 4 (Extensive-form level- k thinking refines outcomes of normal-form level- k thinking) *Consider any finite extensive-form game with perfect recall. Let $\bar{\beta}^1$ be a profile of full-support belief systems and β^1 a profile of full-support beliefs in the associated normal-form consistent with $\bar{\beta}^1$. Then $Z(\bar{L}^k(\bar{\beta}^1)) \subseteq Z(L^k(\beta^1))$ for all $k \geq 1$.*

The proof proceeds by induction. The result for the base-case, level-1, is implied by Proposition 3. At any higher level- k , we focus on players whose information set is reached along the path to the outcome noting that for other players we can simply select an extensive-form level- $(k - 1)$ thinking strategy without affecting the outcome. For any player along the path, we select the first information set, which is well-defined since the game as perfect recall. The extensive-form level- k rational strategy is also rational at this information set with a belief that - as we show - we can confine to normal-form level- $(k - 1)$ strategies of opponents. The result now follows.

The proposition implies in particular that, if initial beliefs are uniform, as often assumed in applications, then for any level k , extensive-form level- k thinking refines the set of outcomes that can be reached by normal-form level- k thinking.

We note that Example 7 and the game used in Subsection 4.1 show that for some games the outcome-refinement of extensive-form level- k thinking is strict at some levels.

3.2 Extensive-Form Level- k Thinking versus Extensive-Form Rationalizability

In Section 2.1 we observed that any normal-form level- k thinking strategy is also level- k rationalizable. In this section, we show that this is not necessarily the case anymore when we consider extensive-form level- k thinking and level- k extensive-form rationalizable strategies. It is not due to a defect in the definitions. Rather, it is due to different

updates conditional on information sets that may occur with and without restrictions on first-order belief systems of extensive-form level- k thinking.

The notion of extensive-form rationalizability is also due to Pearce (1984). However, in contrast its well-known counterpart for normal-form games, there is no treatment of it in standard textbooks on game theory. Consequently, it is much less known although undeservingly so. We follow Battigalli (1997) in allowing for correlated beliefs over opponents' strategies and define it as a reduction procedure on beliefs that subsequently implies a reduction procedure on strategies. The following definition is due to Battigalli (1997) who proves it to be equivalent to Pearce's definition allowing for correlated beliefs.

Definition 4 (Extensive-Form Rationalizability) *Define inductively for every player $i \in N$,*

$$\begin{aligned}
\bar{B}_i^1 & \text{ is the set of player } i \text{'s belief systems.} \\
\bar{R}_i^1 & = \left\{ s_i \in S_i : \begin{array}{l} \text{There exists } \bar{\beta}_i \in \bar{B}_i^1 \text{ with which for every information} \\ \text{set } I_i \in \mathcal{I}_i, s_i \text{ is sequentially rational at } I_i. \end{array} \right\} \\
& \vdots \\
\bar{B}_i^k & = \left\{ \bar{\beta}_i \in \bar{B}_i^{k-1} : \begin{array}{l} \text{For every information set } I_i, \text{ if there exists some profile of} \\ \text{other players' strategies } s_{-i} \in \bar{R}_{-i}^{k-1} \text{ such that } s_{-i} \text{ reaches } I_i, \\ \text{then } \bar{\beta}_i(I_i) \text{ assigns probability 1 to } \bar{R}_{-i}^{k-1}. \end{array} \right\} \\
\bar{R}_i^k & = \left\{ s_i \in S_i : \begin{array}{l} \text{There exists } \bar{\beta}_i \in \bar{B}_i^k \text{ with which for every information} \\ \text{set } I_i \in \mathcal{I}_i, s_i \text{ is sequentially rational at } I_i. \end{array} \right\}
\end{aligned}$$

The set of extensive-form rationalizable strategies is

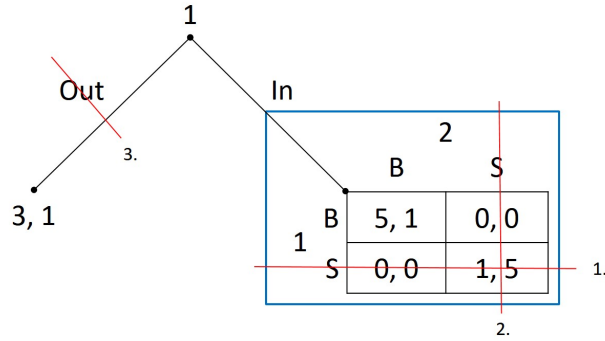
$$\bar{R}_i^\infty = \bigcap_{k=1}^{\infty} \bar{R}_i^k.$$

Battigalli and Siniscalchi (2002) characterize extensive-form rationalizability by common strong belief in rationality. Moreover, for every finite level- k , the k -level extensive-form rationalizable strategies are characterized by k -level mutual strong belief in rationality. Similar to the equivalence between rationalizability and iterated elimination of strictly dominated actions, extensive-form rationalizability is level-by-level strategy-equivalent to iterated elimination of conditionally strictly dominated strategies (Shimoji and Watson, 1998). For each information set of a player, consider the the sub-space of strategy profiles that reach this information set. This is a normal-form information set in the sense of Mailath, Samuelson, and Swinkels (1993). A strategy is conditionally strictly

dominated if there exists a possibly mixed strategy that conditional on such a normal-form information set yields a strictly higher expected utility no matter what strategy profiles in the normal-form information sets are played by opponents.

In the figures, we indicate in the associated normal-form the normal-form information sets associated to information sets in the extensive-form by blue boxes. We also indicate the order of elimination of conditionally strictly dominated strategies (and hence the order of elimination of non-rationalizable strategies) by red lines with numbers attached that represent the level at which the strategy is eliminated. These examples illustrate already some findings. In the Reny game (Figure 1), extensive-form level- k thinking strategies refine level- k extensive-form strategies for $k \geq 1$. In the HMS game (Figure 2), they refine strategies of player 2 at level 1 but otherwise yield the same strategies as higher levels. The extensive-form level- k thinking strategies (with uniform level-1 belief

Figure 3: Battle-of-the-Sexes with an Outside Option I



systems) and k -level extensive-form rationalizable strategies are printed level-by-level in Table 2. (Again, we let “*” stand for any action.)

One of the most prominent examples to demonstrate the forward induction power of extensive-form rationalizability is the battle-of-the-sexes game with an outside option. We use this example to discuss the relationship between extensive-form level- k thinking and level- k extensive-form rationalizability. In particular, we show that in contrast to the analogous normal-form solution concepts, extensive-form level- k thinking does not refine level- k extensive-form rationalizability.

Example 4 (Battle-of-the-sexes with an outside option I) Consider the game in Figure 3. Player 1 moves first, deciding between *Out* and *In*. When he chooses *In*, the battle-of-the-sexes game is played. Both players can choose between *B* and *S*. Player 1

Table 2: Solutions to Outside Option Games

BoS I game	Extensive-form level- k (uniform level-1 belief system)		k -level extensive-form rationalizability	
Level	Player 1	Player 2	Player 1	Player 2
1	$\{(Out, *)\}$	S	$\{(Out, *), (In, B)\}$	$\{B, S\}$
2	$\{(Out, *)\}$	S	$\{(Out, *), (In, B)\}$	$\{B\}$
3	\vdots	\vdots	$\{(In, B)\}$	\vdots

BoS II game	Extensive-form level- k (uniform level-1 belief system)		k -level extensive-form rationalizability	
Level	Player 1	Player 2	Player 1	Player 2
1	$\{(In, B)\}$	S	$\{(Out, *), (In, B)\}$	$\{B, S\}$
2	$\{(Out, *)\}$	B	$\{(Out, *), (In, B)\}$	$\{B\}$
3	$\{(In, B)\}$	B	$\{(In, B)\}$	\vdots
4	\vdots	\vdots	\vdots	\vdots

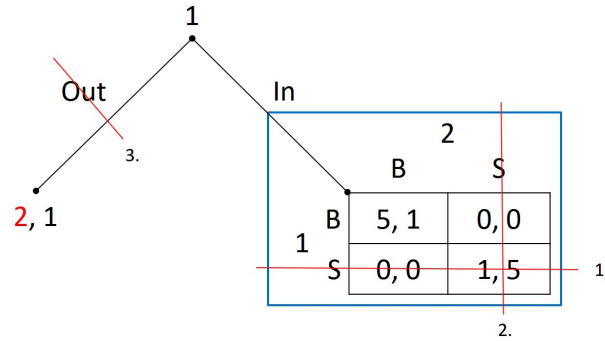
strictly prefers $((In, B), B)$ over $((Out, *), *)$ over $((In, S), S)$ and any other outcome. While for levels $k \leq 2$, extensive-form level- k thinking with uniform initial belief refines level- k extensive-form rationalizable outcomes, at level 3 and higher the extensive-form level- k thinking outcome is distinct from the level- k extensive-form rationalizable outcome. For player 1 at level 1, $(Out, *)$ is rational to a uniform belief over player 2's strategies.⁷ For extensive-form rationalizability, also (In, B) is rational at level 1. Consequently, if player 2's information set is reached at level 2, she knows that player 1 continues with action B to which B is the unique best response by player 2. In contrast, extensive-form level-1 thinking strategies (with uniform initial belief systems) do not reach player 2's information set. Hence, when she is called to play, she must believe that player 1 is level-0, i.e., playing uniformly. With such a belief, S is uniquely rational for player 2. At level 2, player 1 knows that player 2 would play S upon moving In . Given this belief, $(Out, *)$ continues to be rational at the level 2. That is, the prediction of extensive-form level-3 is to observe $(Out, *)$. In contrast, for extensive-form rationalizability, player 1 understands at level 3 that once she moves In , player 2 will select B . Thus, playing B upon moving In is the unique best response of player 1. To summarize, extensive-form level-3 thinking strategies with uniform initial belief yield $((Out, *), S)$ while 3-level

⁷Also Balkenborg and Nagel (2016, p. 398) present such an argument for moving Out at the beginning of the game.

extensive-form rationalizability strategies yield a different outcome, $((In, B), B)$. This example demonstrates that generally extensive-form level- k thinking with uniform belief systems is neither a refinement nor is refined by k -level extensive-form rationalizability.

The outcome of extensive-form level- k thinking in the prior example is driven by the fact that with player 1's uniform belief over player 2's actions, $(Out, *)$ is rational, i.e., yielding a utility of 3, which is larger than 2.5, the expected utility from playing (In, B) against an uniformly mixing player 2. This observation illustrates that outcomes of extensive-form level- k thinking can be sensitive to misspecifications of utilities, which poses interesting issues in experiments. Obviously, the non-robustness is problematic because we can only easily experimentally control payoffs but not necessarily utilities. On the hand, small changes in the utility matter behaviorally in a way that is not picked up by k -level extensive-form rationalizability but is picked up by extensive-form level- k thinking because of restrictions on first-level belief systems. This makes extensive-form level- k thinking interesting in experiments. The prediction of extensive-form level- k thinking can change radically when slightly changing player 1's utility of the outside option in a way that does not affect predictions of extensive-form rationalizability. This is demonstrated in the next example.

Figure 4: Battle-of-the-Sexes with an Outside Option II



Example 5 (Battle-of-the-sexes with an outside option II) Consider the game in Figure 4. This game is identical to the game in Figure 3 except that the outside option yields now a utility of 2 instead of 3 to player 1. The predictions of k -level extensive-form rationalizability remain unchanged and are level-by-level identical to the ones for the game in Figure 3. Yet, extensive-form level- k thinking strategies (with uniform initial

belief systems) of player 1 differ considerably from the ones in Figure 3 (see Table 2). For player 1 at level 1, (In, B) is uniquely sequentially rational to uniform beliefs over player 2's strategies. Player 2 plays optimally S against a uniformly mixing player 1. At the second level, player 1 anticipates this and $(Out, *)$ is uniquely sequentially rational with such a belief. For player 2, she finds B the only rational action at level 2 because she knows by now that once her information set is reached, player 1 continues to play B . Anticipating this at level 3, player 1 selects (In, B) . Player 2 continues with B at level 3 because extensive-form level-2 thinking with uniform initial belief systems of player 1 prescribes $(Out, *)$. That is, her information set should not be reached with such a strategy. Upon reaching the information set nevertheless, she must believe player 1 is extensive-form level-1, i.e., playing (In, B) , and consequently B is the unique best response. Thus, extensive-form level- k thinking yields the extensive-form rationalizable outcome and strategies for $k \geq 3$. But the point of the example is to demonstrate that changing the utility of the outside option to player 1 from 3 to 2 alters dramatically extensive-form level- k thinking strategies. More generally, given uniform initial beliefs of player 1, the prediction flips when player 1's utility of the outside option crosses 2.5. In contrast, this change does not affect the prediction of extensive-form rationalizability.

To sum up, extensive-form level- k thinking with uniform initial belief systems and extensive-form rationalizability are unrelated solution concepts in terms of outcomes, which is in contrast to their normal-form counterparts (see Section 2.1). This is due to restrictions placed on first-level belief systems (i.e., level-0 behavior) for extensive-form level- k thinking. Extensive-form level- k thinking is also very sensitive to misspecifications of utilities.

3.3 Extensive-form Level- k Thinking versus Δ -Rationalizability

In the prior subsection, we argued that extensive-form level- k thinking differs from extensive-form rationalizability due to restrictions placed on first-level belief systems (aka level-0 behavior). Yet, in the literature there are already versions of extensive-form rationalizability that incorporate restrictions on first-level beliefs under the name of (strong) Δ -rationalizability (see Battigalli, 2003, Battigalli and Siniscalchi, 2003, Battigalli and Prestipino, 2013) and extensive-form best response sets (Battigalli and Friedenberg, 2012). Here we state the definition of Δ -rationalizability in a form that facilitates the comparison with extensive-form rationalizability and extensive-form level- k thinking.

Definition 5 (Strong Δ -Rationalizability) For each player $i \in N$, fix a nonempty (measurable) set of restrictions on belief systems $\Delta_i \subseteq \bar{B}_i$. Let $\Delta := (\Delta_i)_{i \in N}$. Define inductively for every player $i \in N$,

$$\begin{aligned}
\bar{B}_i^1(\Delta) & \text{ is the set of player } i\text{'s restricted belief systems } \Delta_i. \\
\bar{R}_i^1(\Delta) & = \left\{ s_i \in S_i : \begin{array}{l} \text{There exists } \bar{\beta}_i \in \bar{B}_i^1(\Delta) \text{ with which for every information} \\ \text{set } I_i \in \mathcal{I}_i, s_i \text{ is sequentially rational at } I_i. \end{array} \right\} \\
& \vdots \\
\bar{B}_i^k(\Delta) & = \left\{ \bar{\beta}_i \in \Delta_i : \begin{array}{l} \text{For every information set } I_i, \text{ if there exists some profile of} \\ \text{other players' strategies } s_{-i} \in \bar{R}_{-i}^{k-1}(\Delta) \text{ such that } s_{-i} \text{ reaches } I_i, \\ \text{then } \bar{\beta}_i(I_i) \text{ assigns probability 1 to } \bar{R}_{-i}^{k-1}(\Delta). \end{array} \right\} \\
\bar{R}_i^k(\Delta) & = \left\{ s_i \in \bar{R}_i^{k-1}(\Delta) : \begin{array}{l} \text{There exists } \bar{\beta}_i \in \bar{B}_i^k(\Delta) \text{ with which for every information} \\ \text{set } I_i \in \mathcal{I}_i, s_i \text{ is sequentially rational at } I_i. \end{array} \right\}
\end{aligned}$$

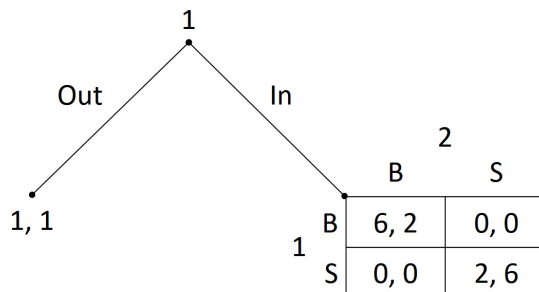
The set of strong Δ -rationalizable strategies is

$$\bar{R}_i^\infty(\Delta) = \bigcap_{k=1}^{\infty} \bar{R}_i^k(\Delta).$$

First note that compared to extensive-form level- k thinking, strong Δ -rationalizability is a reduction procedure on strategies. Yet, it is also easy to see that the set of Δ -rationalizable strategies may be empty. This is naturally the case when the restrictions Δ_i clash with the requirement that beliefs must assign probability 1 to $\bar{R}_{-i}^{k-1}(\Delta)$ when latter set is not ruled out. That is, the crux is in the definition of $\bar{B}_i^k(\Delta)$. We require $\bar{\beta}_i \in \Delta_i$ which might be inconsistent with $\bar{\beta}_i(I_i)(\bar{R}_{-i}^{k-1}(\Delta)) = 1$ and the existence of a strategy profile $s_{-i} \in \bar{R}_{-i}^{k-1}(\Delta)$ that reaches I_i . There are restrictions where the set of Δ -rationalizable strategies is nonempty (see Battigalli, 2003, for non-trivial applications). For instance, in the case of no restrictions, strong Δ -rationalizability is equivalent to extensive-form rationalizability and hence nonempty for every finite game. Most relevant for our comparison with extensive-form level- k , Δ -rationalizability is typically empty if Δ_i is the set of full support belief systems or, even more special, the belief system of uniform beliefs (as often assumed in the level- k literature). We illustrate this in the next example.

Example 6 (Battle-of-the-sexes with an outside option III) Consider the game in Figure 5, a version of a game used in experiments by Cooper et al. (1993). The outside option yields now a payoff strictly lower than the pure equilibrium payoffs in the battle-of-the-sexes game. Consider $\Delta = (\Delta_1, \Delta_2)$ where Δ_1 consists just of the uniform

Figure 5: Battle-of-the-Sexes with an Outside Option III



belief system, i.e., player 1's belief at the beginning of the game and after moving *In* assigns probability $\frac{1}{2}$ to each action, *B* and *S*, of player 2. Similarly, Δ_2 consist of the uniform belief $(\frac{1}{2}, \frac{1}{2})$ of player 2 over player 1's actions *B* and *S* after observing *In*. The strong Δ -rationalizable strategies of the players 1 and 2 at level 1 are (In, B) for player 1 and *S* for player 2. Now, at level 2, neither player can believe in the level-1 Δ -rationalizable strategy of the opponent and have uniform beliefs. Thus, level-2 Δ -rationalizable strategies with Δ being the uniform belief restriction must be empty and so for k -level Δ -rationalizable strategies, for any $k \geq 2$; see Table 3.

Clearly, as the example illustrates, in the case of uniform belief restrictions, we need to give up this restriction at level-2 and higher. This motivates a modification of strong Δ -rationalizability in which we replace $\bar{\beta}_i \in \Delta_i$ with $\bar{\beta}_i \in \bar{B}_i$ in the specification of $\bar{B}_i^k(\Delta)$ of Definition 5. Returning to Example 6, we notice now that the “modified” Δ -rationalizable strategies are nonempty. In particular, they coincide with the first level Δ -rationalizable strategies, (In, B) and *S* for players 1 and 2, respectively. Yet, this is a very strange “solution”. If player 1 realizes that player 2 plays *S*, why wouldn't he best respond with (In, S) . Similarly, if player 2 realizes that player 1 plays (In, B) , why wouldn't she best respond with *B*. The reason is that we require nestedness of strategies in the specification of $R_i^k(\Delta)$ of Definition 5, i.e., $R_i^k(\Delta) \subseteq R_i^{k-1}(\Delta)$, for $k > 1$. However, giving up in the modified definition of strong Δ -rationalizability this nestedness yields our definition of extensive-form level- k thinking. We derive the extensive-form level- k thinking strategies with uniform belief systems of Example 6 in Table 3. It illustrates that the non-nestedness of extensive-form level- k thinking may create cycles, similar to what we have already observed for normal-form level- k thinking in Example 1.

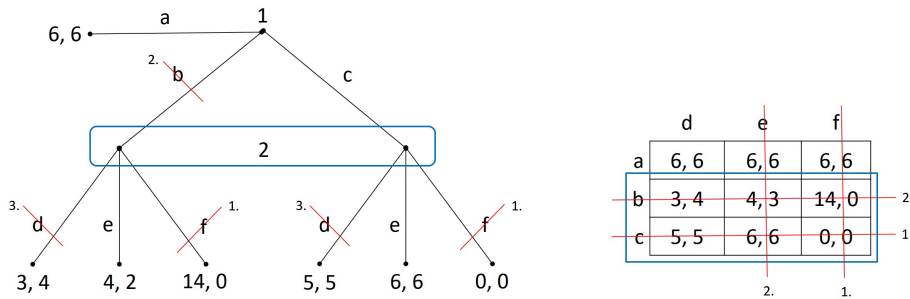
Table 3: Solutions Battle-of-the-Sexes Game III

	Extensive-form level- k (uniform level-1 belief system)		k -level extensive-form rationalizability		k -level Δ rat. (Δ uniform belief system)	
Level	Player 1	Player 2	Player 1	Player 2	Player 1	Player 2
1	(In, B)	S	$(In, *)$	$\{B, S\}$	(In, B)	S
2	(In, S)	B	$(In, *)$	$\{B, S\}$	\emptyset	\emptyset
3	(In, B)	S	$(In, *)$	$\{B, S\}$	\emptyset	\emptyset
\vdots	(altern.)	(altern.)	\vdots	\vdots	\vdots	\vdots

3.4 Extensive-form Level- k Thinking versus Iterated Admissibility

In the battle-of-the-sexes games with an outside option of the previous section, k -level extensive-form rationalizability yields strategies that are equivalent to k -iterative elimination of weakly dominated strategies. Since in the battle-of-the-sexes game with an outside option I, extensive-form level- k thinking with uniform initial beliefs is distinct from k -level extensive-form rationalizability, it demonstrated that extensive-form level- k thinking is also distinct from k -iterative admissibility even if we assume uniform initial belief. Yet, especially with the assumption of initial full-support belief, it is intuitive that sometimes extensive-form level- k thinking retains some features of iterated admissibility that one can think of rationalizability with full-support beliefs. Both feature some form of caution in their beliefs. This is indeed the case in the following example.

Figure 6: HMS2 Game



Example 7 Consider the game of Figure 6 that is a variant of a game discussed in Heifetz, Meier, and Schipper (2021). Let's call it the HMS2 game. In this game,

extensive-form rationalizable strategies are disjoint from strategies remaining from iterated admissibility although in terms of outcomes, iterated admissibility strictly refines the set of extensive-form rationalizable outcomes. (See Table 4, upper part, or the eliminations in Figure 6 (for extensive-form rationalizability) and the associated normal-form (for iterated admissibility). Extensive-form level- k thinking (with uniform initial belief systems) coincides with k -iterated admissibility and is disjoint from k -level extensive-form rationalizability from $k \geq 3$ onward (see the upper part of Table 4). Yet, this is highly sensitive to the utilities. If, for instance, we slightly change player 2's utility of (b, e) from 2 to 3 (let's call it the HMS3 game), then extensive-form rationalizability and iterated admissibility remain unchanged at all levels but now extensive-form level- k thinking (again with uniform initial beliefs) coincides with k -level extensive-form rationalizability and is disjoint from k -iterated admissibility from $k \geq 3$ onward. See the lower part of Table 4. Again, this illustrates the sensitivity of predictions of extensive-form level- k thinking to small changes of utilities. It is due to the assumption of initial uniform beliefs.

Table 4: Solutions to HMS2 and HMS3 Games

HMS2 game	Extensive-form level- k (uniform level-1 belief system)		k -level extensive-form rationalizability		k -iterated admissibility	
Level	Player 1	Player 2	Player 1	Player 2	Player 1	Player 2
1	$\{b\}$	$\{d\}$	$\{a, b, c\}$	$\{d, e\}$	$\{a, b\}$	$\{d, e\}$
2	$\{a\}$	$\{d\}$	$\{a, c\}$	$\{d, e\}$	$\{a, b\}$	$\{d\}$
3	$\{a\}$	$\{d\}$	$\{a, c\}$	$\{e\}$	$\{a\}$	$\{d\}$

HMS3 game	Extensive-form level- k (uniform level-1 belief system)		k -level extensive-form rationalizability		k -iterated admissibility	
Level	Player 1	Player 2	Player 1	Player 2	Player 1	Player 2
1	$\{b\}$	$\{d\}$	$\{a, b, c\}$	$\{d, e\}$	$\{a, b\}$	$\{d, e\}$
2	$\{a, c\}$	$\{d\}$	$\{a, c\}$	$\{d, e\}$	$\{a, b\}$	$\{d\}$
3	$\{a, c\}$	$\{e\}$	$\{a, c\}$	$\{e\}$	$\{a\}$	$\{d\}$

In Subsection 3.3, we realized the difficulty of coming up with a extensive-form rationalizability concept that features initial uniform belief systems even though that would arguably be the most relevant comparison to extensive-form level- k thinking with uniform level-1 belief systems. However, the cautiousness captured in level-1 uniform belief systems could be captured more generally in a rationalizability concept with cautiousness or prudence. Although iterated admissibility captures cautiousness already, iterated

admissibility is a solution concept for the normal form. While Kohlberg and Mertens (1986) argue that the (associated) game in normal-form should be sufficient for solutions to games, there is evidence that behaviorally the extensive-form versus normal-form representation of the games makes a difference (e.g., Cooper and Van Huyck, 2003). It would therefore be desirable to also have an extensive-form rationalizability concept that captures cautiousness. Heifetz, Meier, and Schipper (2021) put forward the following definition of prudent rationalizability.⁸

Definition 6 (Prudent Rationalizability) For any $i \in N$, let $\hat{R}_i^0 = S_i$. For $k \geq 1$, define inductively

$$\hat{B}_i^k = \left\{ \bar{\beta}_i \in \bar{B}_i : \begin{array}{l} \text{For every information set } I_i, \text{ if there exists some profile} \\ s_{-i} \in \hat{R}_{-i}^{k-1} \text{ of the other players' strategies such that } s_{-i} \\ \text{reaches } I_i, \text{ then the support of } \beta_i(I_i) \text{ is the set of strategy} \\ \text{profiles } s_{-i} \in \hat{R}_{-i}^{k-1} \text{ that reach } I_i. \end{array} \right\}$$

$$\hat{R}_i^k = \left\{ s_i \in \hat{R}_i^{k-1} : \begin{array}{l} \text{There exists } \beta_i \in \hat{B}_i^k \text{ such that for all } I_i \in \mathcal{I}_i \text{ player } i \\ \text{with strategy } s_i \text{ is sequentially rational at } I_i. \end{array} \right\}$$

The set of prudent rationalizable strategies of player i is

$$\hat{R}_i^\infty = \bigcap_{k=1}^{\infty} \hat{R}_i^k.$$

Note that this solution concept features non-nested sets of beliefs but nested sets of strategies. It is clear that a full-support belief on a smaller opponents' strategy subspace cannot be an element of the full-support beliefs on larger opponents' strategy subspaces. Thus, the set of k -level prudent belief systems cannot be a subset of the set of $k - 1$ level prudent belief systems. Heifetz, Meier, and Schipper (2021) show that it is nonempty for any finite extensive-form game (including extensive-form games with unawareness).

Meier and Schipper (2022) show that prudent rationalizability is level-by-level strategy equivalent to iterated admissibility in the associated normal form (see Shimoji and

⁸It has been applied to partially identify cautious level- k reasoning in experiments by Li and Schipper (2020). It also been applied to games with unawareness including disclosure games (Heifetz, Meier, and Schipper, 2021; see Li and Schipper, 2019, for experiments), electoral campaigning (Schipper and Woo, 2019) and, with additional belief restrictions, to screening problems (Francetich and Schipper, 2022). In Schipper and Woo (2019), the levels of reasoning embodied in prudent rationalizability have been used to model the political reasoning capabilities of voters.

Watson, 1998, and Brandenburger and Friedenberg, 2011, for related results). Thus, Example 7 demonstrates already how extensive-form level- k thinking with level-1 uniform belief systems differs from prudent rationalizability.

3.5 Extensive-Form Level- k Thinking versus Backward Induction

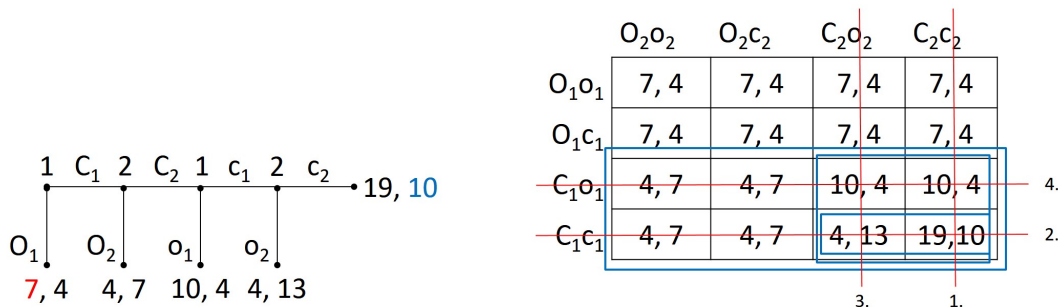
Arguably the most commonly used solution concept to extensive-form games especially with perfect information and non-simultaneous moves is subgame-perfect equilibrium solved by backward induction. Since backward induction is an iterated process, we can consider it also “level-by-level”, i.e., from one subgame to the next of next-higher rank. Consider for simplicity a finite extensive-form game with perfect information, non-simultaneous moves, and a finite horizon. Call the rank of a subgame the maximal number of nodes to reach a terminal node. Backward induction is now defined as follows: At level 1, consider all subgames of rank 1. Select a utility maximizing action of the player that moves in this subgame. Replace the subgame with a terminal node ascribing this utility to this newly created terminal node. Assume we have defined the procedure at level $k - 1$. Then at level k , consider all subgames of rank k (in the original extensive-form game). Select a utility maximizing action of the player that moves in this subgame considering the utilities obtained from the procedure at level $k - 1$. Replace the subgame with a terminal node ascribing this utility to this newly created terminal node. Do this with all subgames of rank k . Since the finite game has a finite horizon, the procedure stops after some finite number of levels.

We can now ask about how extensive-form level- k thinking compares to backward induction up to level k . The following examples demonstrate that backward induction and extensive-form level- k thinking are quite distinct solution concepts.

A prominent game displaying the backward induction logic transparently is the centipede game (Rosenthal, 1981). How does extensive-form level- k thinking compare to backward induction in the centipede game? The example demonstrates that extensive-form level- k thinking may refine level- k backward induction at some levels but be refined by level- k backward induction at some other levels.

Example 8 (Centipede game) *Consider a short version of the centipede game depicted in Figure 7. The set of extensive-form level- k strategies (with uniform initial belief systems) and backward induction strategies for every level k are detailed in Table 1. We note that at level-1, extensive-form level- k thinking (with uniform initial belief system)*

Figure 7: Centipede game



refines the level-1 backward induction strategies and outcomes. This is not surprising as we reap the full power of the assumption on the particular uniform “level-0” behavior while only having the refinement power of backward induction for the game of rank 1, the “tail” and “last leg” of the centipede. At level 2, extensive-form level- k thinking still refines level-2 backward induction but with strategies that differ starkly from extensive-form level-1 thinking as players now take extensive-form level-1 strategies of the opponent into account. At level-3, extensive-form level- k thinking still refines level-3 backward induction outcomes but the strategy of player 2 is now inconsistent with backward induction. The reason is that player 2 does not expect that his second information set is reached. Hence any action at this information set is rational. In contrast, if player 2 follows backward induction and reaches her second information set, she “shrugs her shoulders” and ignorantly continues to do backward induction as if nothing had happened. As from level-4 though, backward induction refines extensive-form level- k strategies (for the same reason that produced the difference at level 3). The outcomes are the same though. To summarize, for levels $k \leq 3$, extensive-form level- k thinking refines level- k backward induction in terms of outcomes. Yet, for levels $k \geq 4$, extensive-form level- k thinking is refined by level- k backward induction in terms of strategies but not outcomes. This changing pattern of the relationship between extensive-form level- k thinking and level- k backward induction highlights the fact that these are conceptually quite different solution concepts although both can be understood as solutions concepts employing some form of inductive elimination of strategies. We note that extensive-form level- k thinking does not display the often criticized “blind belief” in the opponent playing backward induction when an information set is reached that is not on the outcome path.

4 Revisiting Prior Experiments

There is a large literature on testing solution concepts to extensive-form games in experimental game theory. Given our observations with regard to extensive-form level- k reasoning, normal-form level- k reasoning, and k -level extensive-form rationalizability, data sets on games in which forward induction plays a role are of special interest to us. We are very grateful to authors of some previously published experiments for providing us with their data sets. In this section, we report on a simple reanalysis of those extant data sets attempting to glean different aspects of strategic sophistication.

4.1 The Role of Forward Induction Given the Level of Thinking & Uniform First-Level Beliefs

Cooper et al. (1993) conducted experiments on the Battle-of-the-Sexes game with outside options. One of the games used in their experiments is depicted in Figure 8. Although its proper subgame is phrased as an “anti-coordination” game, it is seen easily to be equivalent to a Battle-of-the-Sexes game just by renaming the actions of one player.

Figure 8: Battle-of-the-Sexes Game used by Cooper et al. (1993)

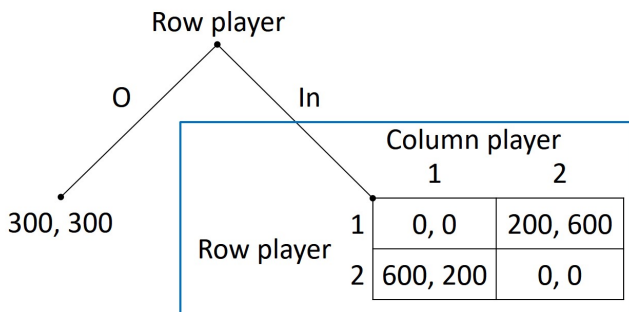


Table 5 presents level-by-level three solutions: extensive-form level- k with uniform initial belief systems, normal-form level- k with uniform initial beliefs, and extensive-form rationalizability. We observe that in this game, extensive-form level- k reasoning is equivalent to k -level extensive-form rationalizability for every level but the first. More importantly, extensive-form level- k thinking is a strategy refinement of normal-form level- k reasoning from level 3 onward and an outcome refinement of normal-form level- k reasoning from level 4 onward. Both, extensive-form level- k reasoning and normal-form level- k reasoning feature levels of reasoning. Yet, extensive-form level- k reasoning also features

Table 5: Solutions to the Battle-the-Sexes Game used by Cooper et al. (1993)

	Extensive-form level- k (uniform level-1 belief syst.)		Normal-form level- k (uniform level-1 belief)		k -level extensive-form rationalizability	
Level	Row player	Column pl.	Row player	Column pl.	Row player	Column pl.
1	$\{(In, 2), (O, *)\}$	2	$\{(In, 2), (O, *)\}$	2	$\{(In, 2), (O, *)\}$	$\{1, 2\}$
2	(O, *)	1	(O, *)	1	$\{(In, 2), (O, *)\}$	1
3	(In,2)	1	(In,2)	$\{1, 2\}$	(In, 2)	1
4	(In,2)	1	$\{(In, 2), (O, *)\}$	1	(In,2)	1
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots

a second dimension of strategic sophistication, namely the ability to update beliefs about opponent’s future behavior given the opponent’s past behavior. Games like the present one, in which extensive-form level- k thinking strictly refines normal-form level- k thinking, provide us with an opportunity to assess the importance of the forward-induction ability beyond just level- k thinking. This becomes apparent at level 3. Under normal-form level- k reasoning for $k = 3$, the column player is indifferent between actions 1 and 2 thinking that the row player chooses Out anyway. In contrast, when the column player’s information set is reached and she gets to play, she now knows under extensive-form level- k reasoning with $k = 3$ that the row player cannot be an extensive-form level-2 reasoner because such a row player would move Out. At extensive-form level-3, the only way for the column player to make sense of the row player’s action to move In is to attribute level-1 to the row player. Here we see the “best rationalizability principle” embodied in extensive-form level- k thinking at work. Rather than believing that the row player is irrational by choosing In, the column player attributes the highest level of rationality consistent with reaching the subgame to the row player (and below her own level-3), which is extensive-form level-1. This makes her realize that the row player plans to play 2 since he moved In already. Thus, she best responds with taking action 1.

Participants played the game for 22 periods and alternated between the row and column player positions. Players were anonymously re-matched. Results of the first 11 periods differed significantly from the last 11 periods and Cooper et al. (1993) reported only on the last 11 periods. Other treatments of the experiment involved variants of the Battle-of-the-Sexes games with outside options. We focus on the game of Figure 8 because extensive-form level- k is a *strict* refinement of normal-form level- k reasoning.⁹

⁹For instance, Cooper et al. (1993) also report on a treatment involving a game similar to Example 6. However, in this game, extensive-form level- k coincides with normal-form level- k and extensive-form rationalizability provides only a coarse solution. Therefore, we do not think we can learn much for our purposes from that treatment w.r.t. the solution concepts discussed here and omit a reanalysis.

See Cooper et al. (1993) for further details of the experimental design.¹⁰

Table 6 describes the percentage of choices consistent with the various solution concepts. First, we observe that a large percentage of choices are consistent with all three

Table 6: Choices in Cooper et al. (1993) Consistent with Solutions

	Extensive-form level- k (uniform level-1 belief system)		Normal-form level- k (uniform level-1 belief)		k -level extensive-form rationalizability	
Level	Row player	Column pl.	Row player	Column pl.	Row player	Column pl.
1	98%	8%	98%	8%	98%	100%
2	20%	92%	20%	92%	98%	92%
3	78%	92%	78%	100%	78%	92%
4	78%	92%	98%	92%	78%	92%

solution concepts. Second and more importantly, we note that apparently the second dimension of strategic sophistication, forward-induction, is just missing in 8% of the column players *at comparable levels of reasoning* (i.e., at level-3, the relevant level for forward induction) and *comparable assumptions on first-level beliefs/level-0 behavior* (i.e., uniform). Normal-form level- k with uniform initial beliefs trivially fits 100% of the data on the column player for $k = 3$ while extensive-form level- k with uniform initial belief systems fits just 8% less at $k = 3$.

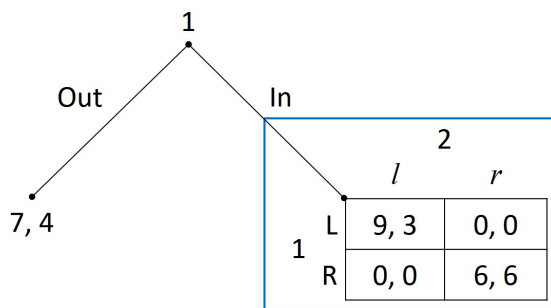
The exercise also offers a glimpse of how extensive-form level- k thinking might be used in experimental game theory. The goal is not so much in winning a horse race among solution concepts in a fitting exercise. Rather, by comparing different solution concepts that differ, very much in the spirit of comparative statics, just in one particular feature of strategic sophistication from each other but are otherwise comparable, we might learn about the prevalence of this feature without interference by other varying features. When comparing the fit of extensive-form level- k with uniform initial belief systems and normal-form level- k with uniform initial beliefs in the experiment, we learn about the prevalence of forward induction *given* comparable levels of thinking and comparable assumptions on initial beliefs.

¹⁰Although we received the data from Cooper et al. (1993), we were not able to fully comprehend them yet given that they were collected more than 30 years ago. Thus, our analysis makes use of the frequencies reported in Cooper et al. (1993, Table 4).

4.2 The Role of Uniform Level-1 Beliefs Given Forward Induction & the Level of Reasoning

Balkenborg and Nagel (2016) study a variant of the battle-of-the-sexes game with an outside option on which nature moves first and selects between the outside option game or the battle-of-the-sexes game without the outside option. We focus here on the subgame consisting of the battle-of-the-sexes game with the outside option as depicted in Figure 9.¹¹ While such a focus on a subgame only is generally problematic when studying solution concepts with forward induction, because beliefs of players might be affected by what happened before the subgame, we do not think that it does affect our analysis of this particular game in a relevant way.

Figure 9: Game used by Balkenborg and Nagel (2016)



The predictions of extensive-form level- k thinking (with uniform initial belief systems) and extensive-form rationalizability are given in Table 7. It is well-known that the outcome of extensive-form rationalizability is equivalent in this game to the prediction by iterated admissibility and strategic stability. Balkenborg and Nagel (2016) refer to this outcome simply as the forward induction outcome. Their interest is on testing it against Harsanyi-Selten equilibrium selection based on risk-dominance, the focal point, and strong backward induction. In particular, since profile (R, r) is both risk-dominant and focal due to symmetry of payoffs, $(6, 6)$, backward induction suggests that player 1 chooses out and guarantees himself the larger payoff of 7. This prediction coincides with extensive-form level- k thinking with uniform level-0 belief systems in this game for $k \geq 1$. In fact, already Balkenborg and Nagel (2016, p. 398) note that Out is player 1's best response to the uniform belief over player 2's actions in the battle-of-the-sexes subgame. The fact that extensive-form level- k thinking with uniform level-0 belief systems differs from k -level extensive-form rationalizability in this game allows us to study the effect of

¹¹In their terminology, it is the “left game”.

level-0 uniform beliefs given forward induction and given comparable levels of thinking.

Table 7: Solutions to the Game of Balkenborg and Nagel (2016)

game	Extensive-form level- k (uniform level-1 belief syst.)		Normal-form level- k (uniform level-1 beliefs)		k -level extensive-form rationalizability	
	Player 1	Player 2	Player 1	Player 2	Player 1	Player 2
1	$(Out, *)$	r	$(Out, *)$	r	$\{(Out, *), (In, L)\}$	$\{l, r\}$
2	$(Out, *)$	r	$(Out, *)$	$\{l, r\}$	$\{(Out, *), (In, L)\}$	l
3	\vdots	\vdots	$\{(Out, *), (In, L)\}$	$\{l, r\}$	$\{(In, L)\}$	\vdots
4	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots

Note that normal-form level- k reasoning with uniform initial beliefs has “no bite” in this game. This is because at level 3, player 1 can have arbitrary beliefs about both level-2 best responses of player 2. This makes both $(Out, *)$ and (In, L) consistent with normal-form level- k thinking for $k \geq 3$. It just underlines the fact that normal-form solution concepts are not always useful for studying extensive-form games. That’s why in what follows we focus on extensive-form level- k thinking and k -level extensive-form rationalizability.

In their experiments, 154 students participated in 13 independent sessions. In each session, the game was played sequentially for 50 rounds which was followed by one round of play using the strategy method (Selten, 1967). Participants were randomly rematched after rounds but maintained their player role throughout the experiment. Sessions differed by the information feedback but results did not differ so that data of the various sessions have been pooled together.

Table 8: Choices in Balkenborg and Nagel (2016) Consistent with Extensive-form Solution Concepts

game	Extensive-form level- k (uniform level-1 belief system)		k -level extensive-form rationalizability	
	Player 1	Player 2	Player 1	Player 2
1	88%	43%	90.2%	100%
2	88%	43%	90.2%	57%
3	88%	43%	2%	57%

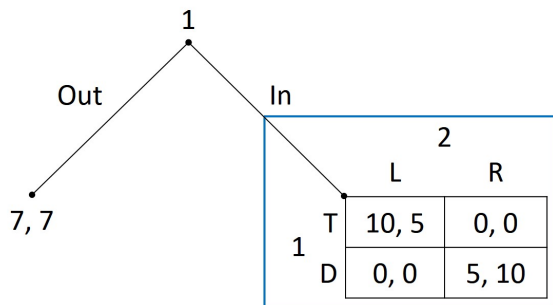
We classify individual choices as in Table 8. Our exercise shows that for the Balkenborg and Nagel (2016) dataset, 88% of the player 1 chooses Out, which is consistent with the prediction of extensive form level- k thinking but not with k -level extensive-form rationalizability from level-3 onward. Since the most substantial difference between k -level extensive-form rationalizability and extensive-form level- k thinking is the assumption of

uniform level-1 belief systems, this difference illustrates the impact of the initial beliefs assumption *given* the level of reasoning (i.e., level 3) and the fact that both solution concepts feature the assumption that players are able to do forward induction.

The picture looks different for player 2. Only 43% of player 2 choose r conditional on the subgame is played. That’s slightly less than for k -level extensive-form rationalizability for $k \geq 2$ (57%). Together this suggests that how participants view the context of the game, as captured by their initial beliefs, may depend on the player role. While the uniform beliefs assumptions seems largely consistent with the behavior of participants in the role of player 1, it is apparently not a descriptive assumption for the majority of participants in the role of player 2.

The last observation suggests further studies. We can use an experiment by Evdokimov and Rustichini (2016) to check for the robustness of the mixed results for player 2. Their experiment makes use of the battle-of-sexes games with outside option depicted in Figure 10. This game has a similar best response structure as the battle-of-the-sexes

Figure 10: Battle-of-the-Sexes Game used by Evdokimov and Rustichini (2016)



game with an outside option I (Example 4) and to the game used by Balkenborg and Nagel (2016). Thus, strategies consistent with various solution concepts are analogous to Table 7 (see Example 4 for more detailed arguments).

There were 230 participants in the experiment. Participants played the game repeatedly. Their player roles could switch between repetitions. Between rounds, they received limited feedback: Player 1 received no feedback about player 2. Player 2 received information on whether or not the Battle-of-the-sexes subgame was reached, i.e., whether or not player 1 moved “In”. There are different treatments that differ in the number of times the game had been repeated, when questions for belief elicitation were asked during the repetitions, and in the incentive structure.¹²

¹²Unfortunately, we do not have the belief elicitation data.

Table 9: Choices in Evdokimov and Rustichini (2016) Consistent with Extensive-form Solutions

game	Extensive-form level- k (with uniform level-1 belief system)		k -level extensive-form rationalizability	
	Player 1	Player 2	Player 1	Player 2
1	62%	22%	98%	100%
2	62%	22%	98%	78%
3	62%	22%	36%	78%

Focusing on the behavioral data from all of their treatments, we classify individual choices as in Table 9. Extensive-form level- k thinking with uniform level-1 belief systems predicts that player 1 always chooses Out at levels above 1, which is consistent with 62% of the choices made by the participants. In terms of player 2’s choice, the extensive-form level- k with uniform level-1 belief systems predicts that player 2 will select R , which is consistent with only 22% of the choices. In contrast, the k -level extensive-form rationalizability fits 78% of the choices made by player 2. However, the extensive-form rationalizability only fits only 36% of the player 1’s choices. While the percentages differ from the corresponding percentages for the Balkenborg-Nagel game in Table 8, the stylized fact from both experiments is that extensive-form level- k with uniform beliefs fits better to the behavior of player 1 while extensive-form rationalizability fits better to player 2. Again, we conclude that while the uniform beliefs assumptions seems more consistent with the behavior of participants in the role of player 1, it is not a descriptive assumption for the majority of participants in the role of player 2. By fitting both solution concepts to the data, we can draw these conclusions *given* comparable levels of reasoning and the ability to do forward induction.

5 Closing Remarks

We extended normal-form level- k thinking to extensive-form games by allowing for updating of beliefs during the play and use of these updated beliefs to make predictions over opponents’ future play. In no way we want to suggest that extensive-form level- k thinking will be the ultimate behavioral solution concept that fits the data on extensive-form games better than other solution concepts. Quite to the contrary, we expect that in abstract choice environments some subjects in experiments may lack the ability to meaningfully draw conclusions from opponents’ past play for opponents’ future play. This ability is like a second dimension of sophistication that is distinct from (but interacts with) the binding cognitive bound. Our hope is that by applying extensive-form level- k

thinking to experimental games and comparing it to normal-form level- k thinking, we can learn about the prevalence of this second dimension of strategic sophistication.

We contrasted extensive-form level- k thinking with other existing iterative solutions concepts to extensive-form games in order to emphasize that there is more than one way to approach levels of thinking in extensive-form games. Unfortunately, extensive-form rationalizability and strong Δ -rationalizability have been understudied in experimental game theory probably because there is no text-book treatment available of these solution concepts. However, similar to Kneeland (2015)'s demonstration that normal-form level- k rationalizability is a valid behavioral solution concept for the empirical study of levels of reasoning, we hope that smart experiments on level- k extensive-form rationalizability will emerge.

The analysis of extensive-form level- k thinking could be taken further in both theoretical and experimental directions. For instance, it might be possible to extend the detailed epistemic analysis of the differences between normal-form level- k thinking and k -level rationalizability by Brandenburger, Friedenberg, and Kneeland (2020) to extensive-form level- k thinking and k -level extensive-form rationalizability. Moreover, while more existing data sets on experimental extensive-form games could be analyzed with extensive-form level- k thinking, we currently think about new experiments with extensive-form games tailor-made for testing extensive-form level- k thinking.

Proofs

Proof of Proposition 1

We prove constructively using induction on the levels.

Base Case: For all $i \in N$, since $\beta_i^1 \in B_i^1$, $L_i^1(\beta^1) \subseteq R_i^1$.

Inductive Hypothesis: For all $i \in N$ and $1 \leq \ell < k$, $L_i^\ell(\beta^1) \subseteq R_i^\ell$.

Inductive Step: We need to show that for every $i \in N$, $L_i^k(\beta^1) \subseteq R_i^k$. Pick any $a_i \in L_i^k(\beta^1)$. By Definition 1, there exists a belief $\beta_i \in \Delta(A_{-i})$ that is certain of $L_{-i}^{k-1}(\beta^1)$ such that a_i is rational with β_i . By the induction hypothesis, such a belief β_i is also certain of R_{-i}^{k-1} . Hence, $a_i \in R_{-i}^k$. \square

Proof of Proposition 2

Since the game is finite, there exists K such that for all $k \geq K$, $R_i^\infty = R_i^k$ for all $i \in N$.

For any $a_i \in R_i^\infty$ there exists $\beta_i \in \Delta(A_{-i})$ such that $\beta_i(R_{-i}^K) = 1$ and a_i is rational for player i given β_i . Since this holds for every player $i \in N$, set $\beta^1 = \beta = (\beta_i)_{i \in N}$. Then $a_i \in L_i^1(\beta^1)$. \square

Proof of Proposition 3

For $k = 1$, note that for any player $i \in N$, if a strategy is rational with respect to a profile of full-support beliefs β_i^1 in the associated normal-form that is consistent with the profile of full-support belief systems $\bar{\beta}^1$, then it is rational with respect to $\bar{\beta}_i^1(I_i)$ conditional on reaching information set I_i . This is due to the fact that conditioning does not alter the relative likelihoods of opponent's strategies reaching the information set. Conversely, if a strategy is rational with respect to a profile of system of full-support beliefs $\bar{\beta}^1$ at *every* information set $I_i \in \mathcal{I}_i$, then is also rational with respect to the profile of full-support beliefs β^1 in the associated normal-form that is consistent with $\bar{\beta}^1$. \square

Proof of Proposition 4

We prove by induction on the levels.

Base Case: By Proposition 3, $\bar{L}_i^1(\bar{\beta}^1) = L_i^1(\beta^1)$ for all $i \in N$. Thus, $\bar{L}^1(\bar{\beta}^1) = L^1(\beta^1)$ and $Z(\bar{L}^1(\bar{\beta}^1)) = Z(L^1(\beta^1))$.

Inductive Hypothesis: For any ℓ with $1 \leq \ell < k$, $Z(\bar{L}^\ell(\bar{\beta}^1)) \subseteq Z(L^\ell(\beta^1))$.

Inductive Step: We need to show $Z(\bar{L}^k(\bar{\beta}^1)) \subseteq Z(L^k(\beta^1))$. Let $z \in Z(\bar{L}^k(\bar{\beta}^1))$.

Let $N(z) \subseteq N$ be the set of players $i \in N$ for whom an information set I_i is reached along the path to z . Since the extensive-form game has perfect recall, each player's set of information sets form an arborescence, i.e., there is a partial order on the set of information sets that orders information sets by precedence. Since it is a partial order, it may have upper bounds, i.e., initial information sets. Yet, along each path, there is a unique upper bound. For any $i \in N(z)$, let I_i denote this first information set of i along the path to z .

Since $z \in Z(\bar{L}^k(\bar{\beta}^1))$, there exists $s \in \bar{L}^k(\bar{\beta}^1)$ with $z(s) = z$. For any $i \in N(z(s))$, there exists $\bar{\beta}_i \in \bar{B}_i^k(\bar{\beta}^1)$ such that $\bar{\beta}_i(I_i)(\bar{L}_{-i}^{k-1}(\bar{\beta}^1)) = 1$ and s_i is sequentially rational at I_i with $\bar{\beta}_i(I_i)$ (with s_i being the i -component of strategy profile s).

For any nonempty $Z' \subseteq Z$, with some slight abuse of notation denote by $S_{-i}(Z') = \{s_{-i} \in S_{-i} : z(s_i, s_{-i}) \in Z', s_i \in S_i\}$. Note that for any nonempty $Z', Z'' \subseteq Z$ with $Z' \subseteq Z''$ implies $S_{-i}(Z') \subseteq S_{-i}(Z'')$. The induction hypothesis, $Z(\bar{L}^{k-1}(\bar{\beta}^1)) \subseteq Z(L^{k-1}(\beta^1))$,

implies $Z(S_i \times \bar{L}_{-i}^{k-1}(\bar{\beta}^1)) \subseteq Z(S_i \times L_{-i}^{k-1}(\beta^1))$. Hence, we have $S_{-i}(Z(S_i \times \bar{L}_{-i}^{k-1}(\bar{\beta}^1))) \subseteq S_{-i}(Z(S_i \times L_{-i}^{k-1}(\beta^1)))$.

Note that for any nonempty $S'_{-i} \subseteq S_{-i}$, $S_{-i}(Z(S_i \times S'_{-i})) \supseteq S'_{-i}$. Thus, $\bar{\beta}_i(I_i)(\bar{L}_{-i}^{k-1}(\bar{\beta}^1)) = 1$ implies $\bar{\beta}_i(I_i)(S_{-i}(Z(S_i \times \bar{L}_{-i}^{k-1}(\bar{\beta}^1)))) = 1$ and $\bar{\beta}_i(I_i)(S_{-i}(Z(S_i \times L_{-i}^{k-1}(\beta^1)))) = 1$.

Define $\beta_i = \bar{\beta}_i(I_i)$. Note that opponents strategies in $S_{-i}(Z(S_i \times L_{-i}^{k-1}(\beta^1))) \setminus L_{-i}^{k-1}(\beta^1)$ do not affect terminal histories, i.e., $Z(S_i \times S_{-i}(Z(S_i \times L_{-i}^{k-1}(\beta^1)))) \subseteq Z(S_i \times L_{-i}^{k-1}(\beta^1))$. To see this, consider any $z(\tilde{s}_1, \tilde{s}_{-i}) \in Z(S_i \times S_{-i}(Z(S_i \times L_{-i}^{k-1}(\beta^1))))$ with $\tilde{s}_i \in S_i$ and $\tilde{s}_{-i} \in S_{-i}(Z(S_i \times L_{-i}^{k-1}(\beta^1)))$. Now, since $\tilde{s}_{-i} \in S_{-i}(Z(S_i \times L_{-i}^{k-1}(\beta^1)))$, there exists $\hat{s}_i \in S_i$ such that $z(\hat{s}_i, \tilde{s}_{-i}) \in Z(S_i \times L_{-i}^{k-1}(\beta^1))$. Because this holds for any $\hat{s}_i \in S_i$, we can set $\hat{s}_i = \tilde{s}_i$. Then $z(\tilde{s}_i, \tilde{s}_{-i}) \in Z(S_i \times L_{-i}^{k-1}(\beta^1))$, which is exactly what we needed to show.

We conclude that $\beta_i(\cdot \mid L_{-i}^{k-1}(\beta^1))$ yields the same expected utilities from strategies as β_i , where $\beta_i(\cdot \mid L_{-i}^{k-1}(\beta^1))$ is β_i conditional on $L_{-i}^{k-1}(\beta^1)$. Moreover, $\beta_i(\cdot \mid L_{-i}^{k-1}(\beta^1)) \in B_i^k(\beta^1)$.

Since s_i is rational with $\bar{\beta}_i(I_i)$ at I_i , it is also rational with $\beta_i(\cdot \mid L_{-i}^{k-1}(\beta^1))$. Thus, $s_i \in L_i^k(\beta^1)$.

This holds for all $i \in N(z(s))$. Since strategies of any other player $j \in N \setminus N(z(s))$ do not affect reaching $z(s)$, we can choose any $s'_j \in L_j^k(\beta^1)$. Note that $z((s_i)_{i \in N(z(s))}, (s'_j)_{j \in N \setminus N(z(s))}) = z(s)$. Since $z((s_i)_{i \in N(z(s))}, (s'_j)_{j \in N \setminus N(z(s))}) \in Z(L^k(\beta^1))$, we have $z(s) \in Z(L^k(\beta^1))$. \square

References

- [1] Agranov, M., Potamites, E., Schotter, A., and C. Tergiman (2012). Beliefs and endogenous cognitive levels: An experimental study, *Games and Economic Behavior* 75, 449–463.
- [2] Alaoui, L., Janezic, K.A., and A. Penta (2020). Reasoning about others' reasoning, *Journal of Economic Theory* 189, 105091.
- [3] Alaoui, L. and A. Penta (2016). Endogenous depth of reasoning, *Review of Economic Studies* 83, 1297–1333.
- [4] Arad, A. and A. Rubinstein. (2012). The 11–20 money request game: A level- k reasoning study, *American Economic Review* 102, 3561–3573.

- [5] Balkenborg, D. and R. Nagel (2016). An experiment on forward vs. backward induction: How fairness and level k reasoning matter, *German Economic Review* 17, 378–408.
- [6] Battigalli, P. (2003). Rationalizability in infinite, dynamic games with incomplete information, *Research in Economics* 57, 1–38.
- [7] Battigalli, P. (1997). On rationalizability in extensive games, *Journal of Economic Theory* 74, 40–61.
- [8] Battigalli, P. (1996). Strategic rationality orderings and the best rationalization principle, *Games and Economic Behavior* 13, 178–200.
- [9] Battigalli, P. and A. Friedenberg (2012). Forward induction reasoning revisited, *Theoretical Economics* 7, 57–98.
- [10] Battigalli, P. and A. Prestipino (2013). Transparent restrictions on beliefs and forward-induction reasoning in games with asymmetric information, *B.E. Journal of Theoretical Economics* 13, 79–130.
- [11] Battigalli, P. and M. Siniscalchi (2003). Rationalization and incomplete information, *B.E. Journal of Theoretical Economics* 3, Article 3.
- [12] Battigalli, P. and M. Siniscalchi (2002). Strong belief and forward induction reasoning, *Journal of Economic Theory* 106, 356–391.
- [13] Bernheim, B.D. (1984). Rationalizable strategic behavior, *Econometrica* 52, 1007–1028.
- [14] Brams, S.J. (1975). *Game theory and politics*, New York: Free Press.
- [15] Brandenburger, A. and E. Dekel (1987). Rationalizability and correlated equilibria, *Econometrica* 55, 1391–1402.
- [16] Brandenburger, A. and A. Friedenberg (2011). The relationship between rationality on the matrix and the tree, mimeo.
- [17] Brandenburger, A., Friedenberg, A. and J. Keisler (2008). Admissibility in games, *Econometrica* 76, 307–352.
- [18] Brandenburger, A., Friedenberg, A., and T. Kneeland (2020). Two approaches to iterated reasoning in games, mimeo.

- [19] Camerer, C.F., Ho, T.-H., and J.-K. Chong (2004). A cognitive hierarchy model of games, *Quarterly Journal of Economics* 119, 861–898.
- [20] Carvajal, A. and H. Zhou (2022). On the sophistication of financial investors and the information revealed by prices, University of California, Davis.
- [21] Cooper, R., DeJong, D.V., Forsythe, R., and T.W. Ross (1993). Forward induction in the Battle-of-the-sexes games, *American Economic Review* 83, 1303–1316.
- [22] Cooper, D.J. and J.B. Van Huyck (2003). Evidence on the equivalence of the strategic and extensive form representation of games, *Journal of Economic Theory* 110, 290–308.
- [23] Cournot, A.A. (1838). *Recherches sur les Principes Mathematiques de la Theorie des Richesses*. Paris: Hachette.
- [24] Crawford, V.P., Costa-Gomes, M.A., and N. Iriberry (2013). Structural models of nonequilibrium strategic thinking: theory, evidence, and applications, *Journal of Economic Literature* 51, 5–62.
- [25] Crawford, V. and N. Iriberry (2007). Level- k auctions: Can a nonequilibrium model of strategic thinking explain the winner’s course and overbidding in private-value auctions? *Econometrica* 75, 1721–1770.
- [26] Evdokimov, P. and A. Rustichini (2016). Forward induction: Thinking and behavior, *Journal of Economic Behavior & Organization* 128, 195–208.
- [27] Farquharson, R. (1969). *Theory of voting*, New Haven: Yale University Press.
- [28] Francetich, A. and B.C. Schipper (2022). Rationalizability screening and disclosure under unawareness, University of California, Davis.
- [29] Feng, J. and X. Wang (2019). Dynamic level- k and cognitive hierarchy models of repeated normal-form games: A Note, mimeo.
- [30] Garcia-Pola, B., Iriberry, N., and J. Kovarik (2020). Non-equilibrium play in centipede games, *Games and Economic Behavior* 120, 391–433.
- [31] Heifetz, A., Meier, M., and B.C. Schipper (2021). Prudent rationalizability in generalized extensive-form games with unawareness, *B.E. Journal of Theoretical Economics* 21, 525–556.

- [32] Heifetz, A., Meier, M., and B.C. Schipper (2019). Comprehensive rationalizability, *Games and Economic Behavior* 116, 185–202.
- [33] Ho T.-H. and X. Su (2013). A dynamic level- k model in sequential games, *Management Science* 59, 452–469.
- [34] Kawagoe, T. and H. Takizawa (2012). Level- k analysis of experimental centipede games, *Journal of Economic Behavior & Organization* 82, 548–566.
- [35] Kneeland, T. (2022). Mechanism design with level- k types: Theory and an application to bilateral trade, *Journal of Economic Theory* 201, 105421.
- [36] Kneeland, T. (2015). Identifying higher-order rationality, *Econometrica* 83, 2065–2079.
- [37] Kohlberg, E. and J.-F. Mertens (1986). On the strategic stability of equilibrium, *Econometrica* 54, 1003–1037.
- [38] Li, Y.X. and B.C. Schipper (2020). Strategic reasoning in persuasion games: an experiment, *Games and Economic Behavior* 121, 329–367.
- [39] Li, Y.X. and B.C. Schipper (2019). Disclosure under unawareness: An experiment, University of California, Davis.
- [40] Mailath, G., Samuelson, L., and J. Swinkels (1993). Extensive form reasoning in normal form games, *Econometrica* 61, 273–302.
- [41] Meier, M. and B.C. Schipper (2022). Conditional dominance in games with unawareness, University of California, Davis.
- [42] Morgenstern, O. (1928). *Wirtschaftsprognose: Eine Untersuchung ihrer Voraussetzungen und Möglichkeiten*, Springer.
- [43] Moulin, H. (1979). Dominance solvable voting schemes, *Econometrica* 47, 1337–1351.
- [44] Nagel, R. (1995). Unraveling in guessing games: An experimental study, *American Economic Review* 85, 1313–1326.
- [45] Osborne, M. and A. Rubinstein (1994). *A course in game theory*, MIT Press.
- [46] Pearce, D.G. (1984). Rationalizable strategic behavior and the problem of perfection, *Econometrica* 52, 1029–1050.

- [47] Reny, P. (1992). Backward induction, normal-form perfection, and explicable equilibria, *Econometrica* 60, 627–649.
- [48] Rosenthal, R.W. (1981). Games of perfect information, predatory pricing, and the Chain-Store-Paradox, *Journal of Economic Theory* 25, 92–100.
- [49] Selten, R. (1967). Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments, in: Sauermann, H. (ed.), *Beiträge zur experimentellen Wirtschaftsforschung*, Tübingen: Mohr, 136–168.
- [50] Schipper, B.C. and H.Y. Woo (2019). Political awareness, microtargeting of voters and negative electoral campaigning, *Quarterly Journal of Political Science* 14, 41–88.
- [51] Shimoji, M. and J. Watson (1998). Conditional dominance, rationalizability, and game forms, *Journal of Economic Theory* 83, 161–195.
- [52] Spohn, W. (1982). How to make sense of game theory, in: Stegmüller, W., Balzer, W., and W. Spohn (eds.), *Philosophy of economics*, Springer-Verlag, 239–270.
- [53] Stahl, D.O. and P.W. Wilson (1995). On players’ models of other players: Theory and experimental evidence. *Games and Economic Behavior* 10, 218–254.
- [54] Stahl, D.O. and P.W. Wilson (1994). Experimental evidence on players’ models of other players, *Journal of Economic Behavior & Organization* 25, 309–327.
- [55] Tan, C.C.T. and S.R.d.C. Werlang (1988). The Bayesian foundations of solution concepts in games, *Journal of Economic Theory* 45, 370–391.
- [56] Zhou, H. (2022). Informed speculation with k -level reasoning, *Journal of Economic Theory* 200, 105384.