# The 11-20 Money Request Game: A Level-*k* Reasoning Study

**Ayala Arad**
School of Economics, Tel Aviv University

**Ariel Rubinstein**
School of Economics, Tel Aviv University
Department of Economics, New York University

**This version**: **October 2011**
**First version**: **January 2010**

**Abstract**

We study experimentally a new two-player game: each player requests an amount between 11 and 20 shekels (NIS). He receives the requested amount and if he requests exactly one shekel less than the other player, he receives an additional 20 shekels. Level-*k* reasoning is appealing due to the natural starting point (requesting 20) and the straightforward best-response operation. Nevertheless, almost all subjects exhibit at most three levels of reasoning. Using two variants of the game we demonstrate that the depth of reasoning is not increased by enhancing the attractiveness of the level-0 strategy or by reducing the monetary cost of undercutting the other player.

Keywords: level-*k* thinking, salience, iterative reasoning

We wish to thank Yaniv Ben-Amy and Aviad Shany for research assistance. We are grateful to the Editor and referees of this journal as well as to Anat Bracha, Vincent Crawford, Gideon Keren, Stefan Penczynski and Pedro Rey Biel for their valuable comments. We acknowledge financial support from the Center for Experimental Social Science at New York University, the Israeli Science Foundation (grant no. 259/08) and the Foerder Institute for Economic Research.

## 1. Introduction

The experimental results in one-shot two-player games often differ significantly from the Nash equilibrium predictions. A major task of the experimental game-theoretic literature is to provide better tools for explaining behavior in such situations. Prominent in this literature is the concept of level-*k* reasoning, first introduced in Stahl and Wilson (1994,1995) and Nagel (1995).

A standard k-level model assumes that the population is partitioned into types that differ in their depth of reasoning. A level-0 type is non-strategic and follows a simple decision rule. A level-*k* type ($Lk$), for any $k \geq 1$, behaves as if he best-responds to the belief that the other player is a level $k - 1$ type. Thus, given a particular game, the model is characterized by: (1) an $L0$ behavior, which is the starting point for iterative reasoning, and (2) a distribution of types. The $L0$ behavior is usually taken to be a uniform distribution on the set of strategies but sometimes the specification also takes into account attraction to salient strategies.

Without having some estimation of the empirical distribution of level-*k* types, it is impossible to use the level-*k* approach to predict behavior in economic settings. A typical study of level-*k* reasoning collects experimental data for a particular game and looks for the best fit to the data given some behavioral assumptions. More precisely, it specifies an $L0$ strategy and assumes that each player is a particular level-k type who sometimes makes an error and chooses a strategy randomly according to a pre-specified distribution. Such a study then finds an ex-post distribution of level-*k* types that best explains the results in a statistical sense (the most frequent types are usually $L1$ and $L2$, whereas higher-level types are rare). Thus, a subject's choice might be attributed to a particular level of reasoning even though it may have been the outcome of a completely different decision rule. Furthermore, this research approach has also been applied to games in which the level-*k* procedure is not so natural or in which the assumed $L0$ strategy is not common knowledge.[1]

Several alternative approaches have been taken in order to improve the assessments of the distribution of level-*k* types. One approach identifies a subject's decision rule by examining his pattern of behavior in a variety of games of the same class and assigning him the decision rule type that most closely fits his overall behavior. Examples of such studies include Stahl and Wilson (1995), Costa-Gomes, Crawford and Broseta (2001) and Costa-Gomes and Crawford (2006).

Another approach confronts the observed choices with additional data that hints on the process of choice in order to determine the level-*k* types. For example, Costa-Gomes, Crawford and Broseta (2001) and Costa-Gomes and Crawford (2006) examined data on information search behavior recorded using MouseLab; Bosch-Domenech, Montalvo, Nagel and Satorra (2002) and Arad (2009) used ex-post explanations of the subjects' choices; Burchardi and Penczynski (2010) analyzed subjects' arguments to convince their teammates to follow their advice; and Selten, Abbink, Buchta and Sadrieh (2003) studied subjects' reasoning by having them construct a computer program that is

able to play any 3 × 3 matrix game.

We take a different route. We design a very simple game that naturally triggers level-*k* reasoning and is not likely to induce other types of decision rules. The *k*-level analysis of the game does not suffer from the problem of ambiguity in specifying the L0 behavior and the identification of types through choice is clear-cut (this identification is supported by subjects' ex-post explanations). The remainder of the introduction presents the game and spells out its merits. The game characteristics are compared to those of some prominent games that were used in the *k*-level literature. In the subsequent sections, we study the game experimentally. The game and two of its variants are used to examine whether the absence of high levels of reasoning found in other studies is a result of certain obstacles to activating level-*k* reasoning in the studied games.

Following is the basic version of the game, which we will refer to as the **11-20 game**:

*You and another player are playing a game in which each player requests an amount of money. The amount must be (an integer) between 11 and 20 shekels. Each player will receive the amount he requests. A player will receive an additional amount of 20 shekels if he asks for exactly one shekel less than the other player.*

*What amount of money would you request?*

In what follows we describe the main aspects of the game that make it particularly suitable for studying k-level reasoning:

(i) *The level-0 type specification is intuitively appealing:* The choice of 20 is a natural anchor for an iterative reasoning process. It is the instinctive choice when choosing a sum of money between 11 and 20 shekels (20 is clearly the salient number in this set and "the more money the better").[2] Furthermore, the choice of 20 is not entirely naive: if a player does not want to take any risk or prefers to avoid strategic thinking, he might give up the attempt to win the additional 20 shekels and may simply request the highest certain amount.

(ii) *Best-responding is straightforward:* Given the anchor 20, best-responding to any level-*k* action is very simple and leaves no room for error. The strategy 19 is the best response to 20, 18 is the best response to 19 and so on (though it is less straightforward that the strategy 20 is the best response to 11).

(iii) *Robustness to the level*-0 *specification:* The choice of 19 is the level-1 strategy given a wide range of level-0 specifications. In particular, 19 is the unique best response to any distribution in which 20 is the most probable strategy. Furthermore, specifying the level-0 behavior as a uniform randomization over the strategies (as is commonly done in games that lack a salient strategy) would not alter the level-1 strategy.[3]

(iv) *Plausibility of the types specification*: The behavior of each level-$k$ type is robust to a wide range of beliefs. For example, if the $L2$ type believes that more than $52.5\%$ of the subjects in the experiment (rather than all the subjects) will choose 19, then his optimal response is still 18. Similarly, the optimal choice of the $L3$ type is 17 if he believes that more than 55% of the subjects will choose 18.

(v) *Clear payoffs:* Unlike some other games that also trigger k-level reasoning, the 11-20 game does not call for social preferences. In particular, if a player believes that his opponent will choose 20 (or some other amount), then requesting one shekel less will give him a bonus of 20 shekels, but **not** at the expense of the other player.

(vi) *Using k-level reasoning is very natural*: The game's payoffs are described explicitly by the best-response function, a characteristic that triggers iterative reasoning. It is hard to think of plausible alternative decision rules for this game. In particular, there is no pure-strategy Nash equilibrium and the game lacks dominated strategies. The only other conceivable rules of behavior we could think of are randomly choosing a strategy or arbitrarily guessing the other player's strategy and best-responding to it.

Let us elaborate on the importance of the last point. If a game triggers attractive decision rules other than $k$-level rules, the estimation of the level-$k$ types of reasoning can be distorted in two ways: (1) Based on observed choice only, we might identify an action as the result of a particular depth of reasoning when in fact it may have been the outcome of a different decision rule. (2) The implementation of alternative decision rules can induce a significant proportion of subjects to choose actions that clearly cannot be attributed to level-$k$ reasoning, and it is not possible to assess how many steps of k-level reasoning those subjects would employ in games where these alternative decision rules are less natural candidates.

It is worthwhile comparing the game's characteristics to those of a few well-known games that were used in the $k$-level literature.

In the normal-form Centipede Game (introduced by Rosenthal (1981) and recently studied using the k-level approach by Kawagoe and Takizawa (2009)), it is natural to implement the iterative elimination of dominated strategies procedure, which is very different from the $k$-level iterated best-response process. Another difference is the lack of any pure Nash equilibrium in our game and the existence of a unique pure Nash equilibrium in the centipede game, which might guide some subjects in their choice.

In Basu's (1994) Traveler's Dilemma[4], the iterated best-response process naturally starts from the highest amount and involves undercutting the other player. However, while in our game a player must choose a number that is exactly one below that of the other player in order to receive the bonus, in the Traveler's Dilemma, choosing any number below that of the other player grants the

bonus. For example, if a player believes that the other player is very likely to choose 300 (though not with certainty), a choice of 298 could be considered a best response to his belief. Another major difference is that in the Traveler's Dilemma, a player might avoid undercutting the other player since he would suffer a utility loss from diminishing the other player's payoff. In contrast, a player who undercuts the other player in our game does not diminish the other player's payoff (and does maximize the sum of payoffs). Finally, following the pure Nash equilibrium in the Traveler's Dilemma is an attractive decision rule which is an alternative to level-*k* reasoning. Branass-Garza, Espinosa and Rey-Biel (2011) classify the k-level and other decision rules that subjects use in this game. They found that only 30% of the explanations included undercutting arguments (mostly level-1 and in a few cases level-2).

In the Beauty Contest (introduced by Nagel (1995)), which is the most extensively studied game in the level-*k* literature, both the iterated elimination of dominated actions and being guided by the Nash equilibrium are prominent decision rules. Furthermore, the best-response function, even in the two-player version of the Beauty Contest, is more complicated and sensitive than the one in our game. The standard specification of level-*k* types in the Beauty Contest is based on the extreme assumption that an *Lk* type believes with certainty that the other player is $Lk - 1$. If the *Lk* type believes that the other player is likely to be an $Lk - 1$ type, though not with certainty, then his best response is very different. This feature of the Beauty Contest might complicate the assessment of the distribution of *k*-level types. In contrast, the *Lk* type's best response in our game is robust to a wide range of beliefs.

To conclude, the level-*k* reasoning process in the 11-20 game is both intuitive and unambiguously specified and hence the identification of levels of reasoning is transparent. This leads us to believe that the game can be a useful tool for research on level-*k* reasoning. For example, the game may be used as a platform for measuring correlations between level-*k* reasoning and other characteristics of the players.

Since the obstacles to activating level-*k* reasoning discussed above are essentially eliminated in the 11-20 game, one can expect that level-*k* reasoning will be observed to a greater extent in an experiment of our game and that the depth of reasoning will be greater than in more complicated games. If this is not the case, then it would call for a search for a more fundamental feature of human reasoning which bounds the depth of reasoning.

Let us turn to the experiments of our game.

## 2. The 11-20 game

*Experimental Design*: The subjects consisted of six classes of undergraduate economics students

at Tel Aviv University (four classes of an Intermediate Microeconomics course and two classes of an introductory course, each consisting of 25-50 students). The subjects had not studied game theory prior to the experiment. The students were offered the opportunity to participate in a short experiment with monetary prizes during class time and all of them decided to participate. Subjects were asked to arrange themselves in the classroom so as not to be able to see their classmates' answers. They were asked to refrain from conversing with each other. Three different forms, which correspond to three versions of the game, were randomly distributed among the subjects (one form to each subject). In this section, we report only on the results for the basic version of the game (described in the Introduction), which was played by 108 subjects out of a total of 233. The instructions for each game appeared on the forms (see the Appendix); no additional instructions were provided. After all the subjects had made their decision, they were asked to write an explanation of their choice on the back of the form. They did not know in advance that they would be asked to do so. After all the forms had been collected, each one was randomly matched to another and each subject received his payoff.

*Results:* Table 1 presents the unique symmetric Nash equilibrium distribution (assuming that players maximize the expected monetary payoff) and the actual distribution of choices in the experiment:

| Action | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|
| Equilibrium | | | | | 25% | 25% | 20% | 15% | 10% | 5% |
| Results | 4% | 0% | 3% | 6% | 1% | 6% | 32% | 30% | 12% | 6% |

**Table 1**: **The 11-20 game** (**n=108**)

The subjects' behavior cannot be explained by Nash equilibrium. The strategy 20 was chosen by 6% of the subjects which is almost identical to the equilibrium prediction. However, only 7% chose the strategies 15 and 16 which is well below the equilibrium prediction of 50%. The vast majority of subjects (74%) chose the actions 17-18-19 which correspond to 1-2-3 levels of reasoning, respectively, whereas in equilibrium they should have been chosen by only 45% of the subjects. According to a chi-square test, the experimental distribution is significantly different from the Nash equilibrium ($p < 0.0001$).

Every strategy in this game could (by definition) be considered a level-*k* strategy for some *k*. However, that does not necessarily mean that it actually reflects k-level reasoning. In particular, we believe that the choices of 11-16 are seldom the outcome of 4 to 9 iterations. Some support for this intuition is obtained by analyzing the subjects' ex-post explanations of their strategies.[5] Only 20% of the subjects chose one of the six strategies 11-16 and these choices were frequently described as

guesses. Few explanations explicitly mentioned best-responding to a specific belief without describing the belief's origin. There were no cases in which the choice of an action in the range 11-15 was explained as 5 to 9-level reasoning and only one subject described the strategy 16 as an outcome of fourth-level reasoning. In contrast, almost all the subjects who chose a strategy in the range 17-19 explained their choices as an outcome of k-level reasoning ("I request 18 shekels since many people believe that the majority will request 20 and thus they will request 19"). Such arguments were somewhat less frequent (77%) among subjects who chose 17.[6]

Note that the level of reasoning attributed to a player is the revealed number of steps carried out from the starting point. It is possible that some subjects realized that the iterative process cycles and returned to the starting point of 20 on reaching strategy 11 (this argument appeared in only six explanations). The bottom line is that, despite this understanding, they chose to conduct only up to three steps of reasoning from the starting strategy.

Following the practice in many of the k-level papers, we have estimated three models. The first model consists of seven types of agents: $L0$ (who chooses 20), $L1$, $L2$, $L3$, $L4$, $L5$ and a type who randomizes uniformly (the frequency of this type could also be interpreted as a common "error rate" of the other types). The second model does not include the $L5$ type and the third does not include the types $L4$ and $L5$. Using the likelihood ratio test, we could not reject the hypothesis that $L4$ and $L5$ are absent at the 5% significance level. Thus, the statistical tests imply that, among these models, the model that consists of the types $L0$-$L3$ (and a noise) best explains the findings.[7]

Studies of other games in the literature also found that it is rare to explicitly observe subjects who practice more than three steps of level-$k$ reasoning. However, as was mentioned in the introduction, in each of these games there was some potential obstacle to activating k-level reasoning (such as the attractiveness of other considerations or decision rules, the complexity or sensitivity of the best-response function and the ambiguity concerning the level-0 behavior). Our most striking finding is that despite the absence of those obstacles in our game and the relatively sophisticated population in the experiment, 80% of the choices reflect up to three levels of reasoning only and higher levels were seldom observed.

One might think that we do not observe higher levels of reasoning since the starting point of the iterative process is not sufficiently salient or because the choice of a number lower than 17 involves giving up too many "certain shekels". In the following sections we examine these conjectures using two variations of the game.

**Comment**. In this paper, we have adopted the definition of a level-$k$ type as a player who best-responds to the belief that the other player is most likely to be a level $k-1$ type. Another version of the k-level model is the Cognitive Hierarchy model introduced by Camerer, Ho and

Chong (2004), in which the level-*k* type best-responds to the conditional distribution of lower types and the distribution of types is assumed to be Poisson with a parameter $\lambda$. Estimating the parameter that best fits our data yields the result: $\lambda = 2.36$.[8] In this model, the strategies that are attributed to the different types are not necessarily as before. In particular, the action 17 is not only chosen by this model's level-3 type but also by all higher-level types. We did not find support in our subjects' explanations for the reasoning behind this alternative model. However, this does not exclude the possibility that subjects indeed used cognitive hierarchy reasoning but neglected to mention this in their explanations. Consider, for example, a subject who justifies his choice of 18 by a belief that most other players choose 19. This subject might also believe that a small proportion of players choose 20 but views this as an insignificant detail. Note that such a belief may be consistent with cognitive hierarchy reasoning but is also consistent with our model's level-2 reasoning due to the extended range of beliefs that are allowed in the 11-20 level-*k* model.

## 3. The Cycle Version

The following version of the game was designed to examine whether enhancing the "status" and salience of the *L*0 action affects the depth of level-*k* reasoning:

*You and another* player *are playing a game in which each player requests an amount of money. The amount must be (an integer) between 11 and 20 shekels. Each player will receive the amount of money he requests. A player will receive an additional amount of 20 shekels if:*

*(i) he asks for exactly one shekel less than the other player*

*or*

*(ii) he asks for 20 shekels and the other player asks for 11 shekels.*

*What amount of money would you request?*

This version was assigned randomly to 72 subjects (the experimental design was described in Section 2). Table 2 presents the unique symmetric Nash equilibrium distribution and the results. For ease of comparison, the results of the basic version are presented again.

| Action | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|
| Equilibrium | | | | | 25% | 25% | 20% | 15% | 10% | 5% |
| Cycle version | 1% | 1% | 0% | 1% | 0% | 4% | 10% | 22% | 47% | 13% |
| Basic version | 4% | 0% | 3% | 6% | 1% | 6% | 32% | 30% | 12% | 6% |

**Table 2**: **The cycle version** (**n**=**72**)

The Nash equilibrium of the cycle version is identical to that of the basic version. Again, the experimental results are significantly different from the Nash distribution according to a chi-square test ($p < 0.0001$).

The cycle version differs from the basic version in one crucial detail. In the basic version, a player who chooses 20 cannot receive the bonus whereas in this version he can (if the other player chooses 11). Thus, the choice of 20 is even more attractive and justifiable than before. This might strengthen the role of 20 as a prominent starting point for iterative reasoning. Although our k-level model for this version remains unchanged, the distribution of strategies may be altered.

The experimental distributions are significantly different in the two versions ($p < 0.0001$). The strategy 20 is indeed more popular than in the basic version (13% vs. 6%). One might have expected that the overall use of level-$k$ strategies (for $k > 0$) would increase due to the enhancement of the level-0 strategy. However, the proportion of subjects who used a strategy in the range of 17-19 did not change significantly ($p = 0.43$); it only increased from 74% to 79%. The main difference in the results is that now most of the choices in the range 17-19 were of the strategy 19. And indeed, the difference between the conditional distributions of the three low level-$k$ types was significant ($p < 0.0001$). In other words, enhancing the $L0$ strategy induced subjects who used k-level reasoning to use fewer steps of iterative reasoning. The explanations of the strategies in this range were very similar to those provided in the basic version. Level-$k$ arguments appeared in all the explanations of the subjects who chose 18 and 19 and in 4 out of the 7 explanations of subjects who chose 17.

To summarize: the role of the strategy 20 as the $L0$ strategy was enhanced in this version and indeed the proportion of subjects who chose the $L0$ strategy doubled (though it still remained small). This might have no effect on the distribution of k-level types. We found that the proportion of subjects who chose the $L1$, $L2$ or $L3$ strategies (altogether) remained unchanged. However, the proportion of subjects who chose the $L1$ strategy increased dramatically while the proportion of those who chose $L3$ decreased. It appears that in this version, many potential level-$k$ types recognized the enhanced status of $L0$, expected it to be chosen frequently, and thus chose the $L1$ strategy.

## 4. The Costless Iterations Version

In the basic version of the game, best-responding to 20 requires giving up one "certain" shekel in the attempt to win an additional 20 shekels. In the same manner, performing another iteration and choosing 18 required giving up an additional shekel. One could suspect that the cost of additional iterations in the basic version is why subjects stop at the level-3 strategy (17) and do not perform

another iteration (which would require giving up an additional certain shekel). That is, sacrificing 4 shekels in an attempt to win the 20-shekel bonus feels like too high a price to pay.

The following version of the game was designed to examine this conjecture:

*You and another player are playing a game in which each player chooses an integer in the range 11-20.*

*A player who chooses* 20 *will receive 20 shekels (regardless of the other player's choice).*

*A player who chooses any other number in this range will receive 3 shekels less than in the case where he chooses 20. However, he will receive an additional amount of 20 shekels if he chooses a number that is one less than that chosen by the other player.*

*Which number would you choose?*

In this version of the game, the "cost" of choosing any integer in the range 11-19 is identical: instead of receiving 20 shekels for certain, each of these numbers guarantees only 17 shekels. Therefore, undercutting a strategy within the range 11-19 does not involve an additional sacrifice.

53 subjects played this version of the game. Of these, 27 were from among those who had been randomly assigned to play one of the three versions of the game, as described in Section 2. 26 additional students were from the same pool of students (first-year Economics undergraduates at Tel Aviv University), but were recruited from a separate class. The distributions of strategies in these two groups were very similar and were pooled together.

Table 3 presents the equilibrium prediction and the actual distribution of strategies in this version:

| Action | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|
| Equilibrium | | | | 10% | 15% | 15% | 15% | 15% | 15% | 15% |
| Results | 0% | 4% | 0% | 4% | 4% | 4% | 9% | 21% | 40% | 15% |

**Table 3**: **The costless iterations version** (n=53)

According to a chi-square test, the experimental distribution of choices is significantly different from the Nash equilibrium ($p < 0.0001$), as was the case in the other two versions. A vast majority of the subjects (84%) chose the strategies corresponding to $L0$, $L1$, $L2$ or $L3$. This proportion is not significantly different from that obtained in the cycle version ($p = 0.24$) or in the basic version ($p = 0.5$). The conditional distribution of the $L1$, $L2$ and $L3$ strategies in the costless iterations version is also not significantly different from that obtained in the cycle version ($p = 0.96$) and it is significantly different from that in the basic version ($p = 0.0001$).[9] The similarity of the results here to those in the cycle version is probably due to the framing in the current version which also

emphasizes the level-0 strategy. According to the subjects' explanations, the choices in the range 17-19 (made by 70% of the subjects) generally involved an iterative process, whereas the choices in the range 14-16 (made by eight subjects) were arbitrary guesses.

Thus, it appears that the cost of performing an additional iteration (i.e., losing an additional certain shekel) is not the reason that subjects perform no more than three iterations in the other two versions of the game.

## 6. Conclusion

In this paper, we have presented a new game for the study of level-*k* reasoning. The game is easily understood, the k-level reasoning process is both intuitive and unambiguously specified, and there is no complexity in carrying out the best response. We have found that, despite these properties, subjects did not use more than three steps of reasoning. The distribution of k-level types may depend on the particular population (for example, Chess masters may exhibit higher levels of reasoning). Nevertheless, the results here, obtained from a fairly sophisticated population of subjects, suggest that it would be rare to observe level-4 reasoning or higher in any one-shot game.

This conclusion is in line with previous studies which found almost no traces of higher levels of reasoning in other games. However, as explained earlier, in those games there were some expected obstacles to the activation of level-*k* reasoning. The current game was designed to eliminate those obstacles and thus our finding that higher levels were seldom observed is intriguing.

We also studied two variations of the game and demonstrated that: (1) enhancing the $L0$ strategy does not increase the overall use of level-*k* strategies (for $k > 0$) and lowers the levels used, conditional on using these level-*k* strategies, and (2) subjects do not use more than 3 steps of reasoning even when the monetary cost of using each level of reasoning (for $k > 0$) is identical. These results rule out the possibility that the lack of high levels of reasoning are due to the vagueness concerning the anchor of the iterative process or the monetary cost of performing an iterative step.

We infer that subjects do not continue to level-4 or higher levels since they do not believe that others will use more than two iterative steps. This may be a consequence of fundamental features of human reasoning related to the rare use of high-order intentionality in everyday language.

Note that the statement we attribute to a level-1 type is "I think that [he is doing something]". The statement we attribute to a level-2 type is "I think that [he thinks that I am doing something]", in which the level-1 statement is embedded. Explicitly expressing the considerations of a level-4 type would involve three such embedments. Whether these considerations are implicit or explicit, it seems that behavior that is consistent with such considerations is unnatural.

This may be rooted in a phenomenon observed by psychologists, such as Kinderman, Dunbar and Bentall (1998): most subjects do not understand a sentence such as: “A thinks that B thinks that A thinks that B thinks that A is doing something", which is attributed to our level-4 type. (Almost all subjects do understand sentences that we attribute to lower levels.) They claim that this kind of sentence goes beyond the limit of reasoning normally used in real life and that most everyday situations probably do not require more than second-order intentionality. Literary scholars have also recently argued that "the zone of cognitive comfort" seems to be very limited. Zunshine (2011) claims that using sentences with a higher number of embedments than in the sentence "I know that you think that he wants you to believe that she was angry at him" is rare and appears in the literature as a challenge to the reader.

Note that iterative k-level reasoning requires less cognitive load than does comprehending the above sentence. A level-3 type need not say "I think I am facing a level-2 type, who thinks he’s facing a level-1 type, who believes he’s facing a level-0 type" but rather can simply use the following algorithm: "Since level-0 will play *A*, level-1 will play *B* and therefore level-2 will play *C*. Thus, I should play *D*." At each stage of the algorithm a player needs to remember only the outcome of the last stage rather than the entire meaning of the sentence, as is required in everyday language. Nonetheless, the scarcity of level-4 types in strategic games is consistent with the cognitive limit on the number of embedments observed in everyday language and may be driven by it.

We propose the 11-20 game as a tool for future research on level-*k* reasoning. Since the game is very simple and almost explicitly guides players to use level-k reasoning, we speculate that the distribution of levels of reasoning in more complicated games will be stochastically dominated by the one in this game. Thus, the game can be used to evaluate the plausibility of a distribution of types that best fits the data in other games. (Needless to say, we do not argue that it is impossible to come up with a game that is even more guiding and hence inducing even higher levels of reasoning.) Furthermore, since almost any choice of subjects in this game can be clearly identified as a particular level of reasoning, the game can be used to study correlations between the depth of reasoning and behavior in other games or other personality traits[10] and to explore the effect of various aspects of the environment on the levels of reasoning.

**References**


Arad, Ayala (2009). "The Tennis Coach Problem: A Game-Theoretic and Experimental Study", *The B.E. Journal of Theoretical Economics* (forthcoming).

Basu, Kaushik (1994). "The Traveler's Dilemma: Paradoxes of Rationality in Game Theory". *American Economic Review,* 84, 391-395.

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

Branas-Garza, Pablo, Maria Paz Espinosa and Pedro Rey-Biel (2011). "Travelers' Types". *Journal of Economic Behavior & Organization*, 78, 25-36.

Burchardi, Konrad B. and Stefan P. Penczynski (2010). "Out of Your Mind: Estimating the Level-k Model". Working Paper.

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

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

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

Crawford, Vincent and Nagore Iriberri (2007). "Fatal Attraction: Salience, Naivete, and Sophistication in Experimental Hide-and-Seek Games", *American Economic Review,* 97, 1731-1750.

Georganas, Sotiris., Paul J. Healy, and Roberto Weber. (2010). "On the Persistence of Strategic Sophistication". Working Paper, University of Bonn.

Kawagoe, Toshiji and Hirokazu Takizawa (2009). "Level-k Analysis of Experimental Centipede Games," SSRN (Social Science Research Network) ID:1289514.

Kinderman, Peter, Robin Dunbar, and Richard, P. Bentall (1998). "Theory-of-Mind Deficits and Casual Attributions", *British Journal of Psychology,* 89, 191-204.

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

Penczynski Stefan P. (2011). "Strategic Thinking: The Influence of the Game". Working Paper.

Rosenthal, Robert (1981). "Games of Perfect Information, Predatory Pricing, and the Chain Store," *Journal of Economic Theory,* 25, 92-100.

Rubinstein, Ariel, Amos Tversky and Dana Heller (1996). "Naive Strategies in Competitive

Games". *Understanding Strategic Interaction - Essays in Honor of Reinhard Selten*, Eds. Albers, W. et al, Springer-Verlag, 394-402.

Selten, Reinhard, Klaus Abbink, Joachim Buchta and Abdolkarim Sadrieh. (2003). "How to Play (3 × 3)-Games. A Strategy Method Experiment," *Games and Economic Behavior*, 45, 19-37.

Stahl, Dale O. and Paul W. Wilson (1994). "Experimental Evidence on Players' Models of Other Players". J*ournal of Economic Behavior and Organization*, 25, 309-327.

Stahl, Dale O. and Paul W. Wilson (1995). "On Players' Models of Other Players: Theory and Experimental Evidence", *Games and Economic Behavior*, 10, 213-54.

Zunshein, Lisa (2011). "1700-1775: Theory of Mind, Social Hierarchy, and the Emergence of Narrative Subjectivity." *The Emergence of Mind: Representations of Consciousness in Narrative Discourse in English, 700 – the Present*. Ed. David Herman. Lincoln: University of Nebraska Press.

**Footnotes**

1. An example is Crawford and Iriberri (2007) who analyzed the hide-and-seek game introduced in Rubinstein, Tversky and Heller (1996). Crawford and Iriberri (2007) found that given their assumptions on the level-0 behavior, the distribution of types that best fits the data included many level-3 and level-4 types. However, Penczynski's (2011) analysis of subjects' arguments in the same hide-and-seek game raises doubts as to the interpretation of the best fit as a real distribution of level-*k* types. The arguments reveal that subjects do not have a common starting point for iterated reasoning (i.e., level-0) and that only a few subjects practice more than two steps of reasoning given their own starting point.

2. The number 20 is also the focal point in the range 11-20 in the following sense: if two people play a game in which each chooses a number in the above range and is awarded a prize only if both choose the same number, then clearly they would choose 20.

3. The strategy 19 remains a best response to any belief that does not assign to any action $20 - i$ a probability that is higher by $5i$ % than the probability assigned to the strategy 20.

4. In the Traveler's Dilemma, each of two players requests an amount between $180 and $300 and receives the amount he requests. In addition, a player who makes the lower request is paid $5 by the other player.

5. Subjects' explanations were classified by a research assistant who otherwise was not involved in the research. The research assistant chose the categories himself. The classification was straightforward and we confirmed the appropriateness of his analysis.

6. A typical explanation for the choice of 20 was "The chances of playing against someone who requests one shekel more than me is very low. Therefore, I prefer to receive the highest sum with certainty."

7. In a model that includes the types $L0$, $L1$, $L2$, $L3$ and $L4$, the estimated frequencies are: $L0 = 0.05$, $L1 = 0.13$, $L2 = 0.37$, $L3 = 0.40$ and $L4 = 0.05$ with an error rate of $0.26$.

In a model that includes the types $L0$, $L1$, $L2$ and $L3$, the estimated frequencies are: $L0 = 0.05$, $L1 = 0.13$, $L2 = 0.39$ and $L3 = 0.43$, with an error rate of 0.32.

8. We used the calculator "Cognitive Hierarchy Theory of One-shot Games" constructed by Colin Camerer, Teck-Hua Ho and Juin-Kuan Chong .

9. When we treat the strategies 11-16 as one category (which is required due to the small number of choices in these strategies) and each of the other strategies as a separate category, the distributions in the costless iterations version and the cycle version are not significantly different ($p = 0.77$).

10. A recent example of a study on the persistence of k-level reasoning is Georganas, Healy and Weber (2010). The basic game that they study is a 7x7 matrix game that lacks a simple verbal description. One may find some similarities between their game and the 11-20 game. However, the level-*k* specification in their game is less appealing. In particular, their game lacks a salient strategy that could serve as an anchor, it includes dominated actions and it does not allow the distinction between the fourth-level strategy and playing the pure Nash equilibrium.

## Appendix: Translation of the three forms used in the experiments

### (1) The 11-20 Game

Name: _____

Sex: Male / Female

Your personal details are confidential and will be used only to make the payment.

---

You are randomly matched to play a game against one of the students in this class. In the game, each of you requests an amount of money (an integer) between 11 and 20 shekels. Each participant will receive the amount he requests. A participant will receive an additional 20 shekels if he asks for exactly one shekel less than the other player.

You will receive your payment in the next lesson, without knowing against whom you played.

What amount of money do you request?

### (2) The Cycle Version

Name: _____

Sex: Male / Female

Your personal details are confidential and will be used only to make the payment.

---

You are randomly matched to play a game against one of the students in this class. In the game, each of you requests an amount of money (an integer) between 11 and 20 shekels. Each participant will receive the amount he requests. A participant will receive an additional 20 shekels if:

(i) he asks for exactly one shekel less than the other player,

or

(ii) he asks for 20 shekels and the other player asks for 11 shekels.

You will receive your payment in the next lesson, without knowing against whom you played.

What amount of money do you request?

## (3) The Costless Iterations Version

Name: _____

Sex: Male / Female

Your personal details are confidential and will be used only to make the payment.

---

You are randomly matched to play a game against one of the students in this class. In the game, each of you selects an integer between 11 and 20 shekels.

- If you choose the number 20, you will receive 20 shekels (regardless of the other player's choice).

- If you choose any other number in this range, you will receive 3 shekels less than in the case where you choose 20. However, you will receive an additional amount of 20 shekels if you choose a number that is lower by exactly one than the number chosen by the other player.

You will receive your payment in the next lesson, without knowing against whom you played.

Which number do you choose?