Logical Omniscience at the Laboratory

Ralph C. Bayer, Ludovic Renou

PII: S2214-8043(16)30057-X
DOI: 10.1016/j.socec.2016.06.003
Reference: JBEE 217

To appear in: *Journal of Behavioral and Experimental Economics*

Received date: 19 May 2014
Revised date: 19 May 2016
Accepted date: 6 June 2016

Please cite this article as: Ralph C. Bayer, Ludovic Renou, Logical Omniscience at the Laboratory, *Journal of Behavioral and Experimental Economics* (2016), doi: 10.1016/j.socec.2016.06.003

Highlights

- We experimentally measure how many steps of counterfactual reasoning humans can do.
- In order to separate cognitive ability from other behavioural influence factors in games we use a novel computerized version of the Dirty Faces Game.
- We find that on average our subjects can perform two to three steps of iterative reasoning.
- This shows that subjects are better at iterative reasoning than suggested by studies that do not control for subjects? beliefs about the rationality of others or social preferences.

Logical Omniscience at the Laboratory*

Ralph C. Bayer† & Ludovic Renou‡

University of Adelaide & *University of Essex*

June 10, 2016

Abstract

This paper investigates the ability of individuals to make complex chains of reasoning, similar to those underlying the logic of iterated deletion of dominated strategies. Controlling for other-regarding preferences and beliefs about the rationality of others, we show, in the laboratory, that the ability of individuals to perform complex chains of iterative reasoning is better than previously thought. We conclude this from comparing our results with those from studies that use the same game without controlling for confounding factors. Subjects were able to perform about two to three iterations of reasoning on average as measured by our version of the Red-Hat Puzzle.

Keywords: iterative reasoning, depth of reasoning, logical omniscience, rationality, experiments, other-regarding preferences.

JEL Classification Numbers: C70, C91.

*We thank the Economic Design Network Australia and the University of Adelaide for funding. We are grateful to Karl Schlag, Konstantinos Tatsiramos, and seminar and conference participants at the Australian National University, the University of Melbourne, the University of Western Australia, the California Institute of Technology, and the University of East Anglia.

†Department of Economics, 10 Pulteney Street, Adelaide 5005, Australia. Phone: +61 (0)8 8313 4666, Fax: +61 (0)8 8223 1460. `ralph.bayer@adelaide.edu.au`

‡Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom. `lrenou@essex.ac.uk`

The view that machines cannot give rise to surprises is due, I believe, to a fallacy to which philosophers and mathematicians are particularly subject. This is the assumption that as soon as a fact is presented to a mind all consequences of that fact spring into a mind simultaneously with it. It is a very useful assumption under many circumstances, but one too easily forgets that it is false. A natural consequence of doing so is that one then assumes that there is no virtue in the mere working out of consequences from data and general principles. Alan Turing (1950)

1 Introduction

Logical omniscience and rationality are two central assumptions in Game Theory. A player is *logically omniscient* if he knows *all* logical implications of his knowledge and *rational* if he chooses optimal strategies given his knowledge and beliefs. The aim of this paper is to experimentally measure the degree of logical omniscience (and rationality) of individuals, controlling for other-regarding preferences and beliefs about the rationality and omniscience of others.

All experimental attempts to measure the degree of logical omniscience (and rationality) in humans by analyzing behavior in *strategic games* necessarily conflate auxiliary hypotheses on subjects' perception of the cognitive abilities and preferences of others. Bounded rationality and other factors (strategic uncertainty, social preferences, overconfidence, etc.) cannot be cleanly separated in such experiments. This paper proposes a novel experimental design, which makes it possible to measure the degree of logical omniscience and rationality of individuals with as few confounding factors as possible.

To see that measurement without confound is difficult, consider the seminal beauty contest game (Nagel, 1995). Deducing the level of a subject's level of logical abilities from the number chosen is bound to be biased. For instance, a scholar of game theory would choose a reasonably high number if she believes that the iterative abilities of others are low, despite having the ability to iterate to the equilibrium choice of zero. After all, the optimal choice is to best-reply to one's conjecture about the choices of others, not necessarily to play equilibrium (unless one conjectures that others play according to equilibrium). Therefore, direct measures of logical abilities from observed behavior in strategic-form games are bound to be biased, as they do not take into account that play is not only a result of cognitive abilities, but also of a player's beliefs about the play of others. Agranov et al. (2012) show that the beliefs about the rationality of others indeed play an important role, as in their study the number of iterations performed in the guessing game varies in the expected way, when beliefs about the cognitive abilities of other players are manipulated. Disentangling own cognitive ability and beliefs is made even more difficult by the fact that not all subjects adjust their iteration depth in the same way. Gill and Prowse (2015) show that only subjects with high cognitive abilities adapt their behavior to information about the cognitive abilities of others. Aloaui and Penta (2015) develop a model of how own cognitive abilities and beliefs about the the cognitive abilities of others translate into behavior and find support for their theory using the 11-20 game (Arad and Rubinstein, 2012). Social preferences and preferences for social efficiency are additional confounding factors.

This paper acknowledges the problem of the confound and makes a methodological contribution towards solving it. We offer an experimental design that makes it possible to measure the ability of individuals to perform chains of iterative reasoning with as little confounding factors as possible, without sacrificing a game-like structure. The resulting measure of logical omniscience can be used as an explanatory variable for observed behavior in strategic-form games, which allows for an assessment of the degree to which limited cognitive abilities contribute to deviations from Nash behavior.

The experiment we designed is a variant of the Red Hat Puzzle (also known as the Dirty Faces Game), in which we control for other-regarding preferences and beliefs about the rationality of others. In the Red Hat Puzzle (RHP thereafter), a player has to determine her type (hat color) by the use of iterative reasoning. For this purpose the player can use her knowledge about the types of the other players and the other players' actions. The distribution of types determines how many iteration steps a player has to perform in order to arrive at the correct answer.[1] In its original form (as used by Weber (2001) or Bayer and Chan (2009)), the RHP suffers from the same problems as other interactive games when used to measure subjects' iteration ability. Players have to rely on the iterative abilities of other players. Therefore, not only their own iterative ability matters but also their beliefs about the ability of others, beliefs about beliefs about the ability of others, etc.[2] Social preferences might also play a role. To overcome this problem we do the following: we transform the RHP into an interactive decision problem where every "player" at each move has a unique logically correct answer. In each game, a single human player plays with computer players only.[3] The computer players are programmed to be logically omniscient, i.e. they always choose the logically correct answer. This fact is communicated to the human player. In this setup a human player, who is able to perform the necessary number of iteration steps for a particular puzzle, can

[1]A detailed description of the puzzle will be given below.

[2]The methodology used in this paper has first been described in a conference paper (Bayer and Renou, 2007), which is based on the data from a pilot for this study. The pilot only contained a few sessions of one of the six treatments presented here. The conference paper's purpose was to describe the methodology, while this paper shows how behavior changes across the treatment dimensions.

[3]For others experimental designs with automated opponents, see Johnson et al. (2002) and McKinney and Van Huyck (2007).

fully rely on the other players' logical omniscience. Additionally, we do not have to worry about the influence of social preferences as the human player does not interact with other humans.[4] While this transformation makes the RHP a decision problem, it still remains interactive. Computer-players interact with the human-subjects in that their "actions" will depend on the action of the human subjects, and vice versa. With this procedure, we can cleanly isolate and measure the iteration ability of humans in an interactive situation by varying the type distribution within a subject.

Our experiments highlight two interesting patterns. Firstly, subjects were able to perform about two to three steps of iterative reasoning on average, more than the one to two steps typically measured in similar games without control for beliefs about the rationality of others. It is important to stress that comparisons with previous studies that do not control for social preferences or beliefs about the rationality of others are difficult. Without additional assumptions on the preferences and beliefs about the rationality of others, it is not possible to infer the ability to perform steps of iterative reasoning from observed play in strategic-form games, for example. A second result refers to learning: to our surprise, subjects did not only learn from observation (feedback). Introspection alone was sufficient for subjects to perform better when playing the same puzzles for a second time.[5] Our econometric analysis is organized around these two themes (Section 4).

This paper contributes to the large literature on iterative reasoning in games e.g., McKelvey and Palfrey (1992), Beard and Beil (1994), Nagel (1995), Ho, Camerer and Weigelt (1998), Goeree and Holt (2001), Van Huck, Wildenthal and Battalio (2002), Cabrera, Capra and Gómez (2006), to name just a few.[6] A recurring feature of many of these studies is the use of games solvable by iterated deletion of strictly or weakly dominated strategies.[7] In these studies, the ability of individuals to perform iterative reasoning is associated with their ability to iteratively delete dominated strategies. Centipede games (e.g., McKelvey and Palfrey (1992)) and beauty contest games (introduced to the literature by Nagel (1995)) are two of the most commonly used games in that literature. However, in those games, iterating to the equilibrium might actually not be optimal for a subject. E.g, in a centipede game, a fully rational and omniscient player will pass instead of ending the game right away, as prescribed by a subgame-perfect Nash equilibrium, if she beliefs that the opponent does not understand equilibrium logic and will pass given the next move. Without controlling for beliefs about the rationality and logical omniscience of others, failure to play the equilibrium cannot be

[4]Naturally, we cannot exclude the possibility that the subjects had concerns for the well-being of other persons affected by their decisions, e.g., the experimenters, other students (perhaps because the funding used in the experiment could have helped these students), etc. This is unlikely to play a major role, though. For instance, Frank (1998) and Fleming and Zizzo (2015) have not found evidence of altruism towards experimenters.

[5]A similar observation is made in Weber (2003).

[6]We refer the reader to chapter 5 of Camerer (2003) for a survey of this literature.

[7]Note that the solution concept of iterated deletion of weakly dominated strategies requires more stringent conditions than common knowledge of rationality (see Brandenburger et al. (2008)).

interpreted as limited ability to perform iterated reasoning.[8] Consequently, a researcher interested in the ability of humans to perform chains of iterative reasoning might underestimate the actual ability of humans when relying on choices in beauty contest or centipede games alone. The same is true, to our knowledge, for all studies of interactive games aiming to measure the iteration abilities of humans.[9]

Two closely related experimental studies are Weber (2001) and Bayer and Chan (2009). Weber implements the red hat puzzle as a dynamic game of incomplete information between two or three human players. Bayer and Chan replicate Weber's experiment and compare the replication with a modified version of Weber's game, where equilibria in weakly dominated strategies are eliminated.[10] As already argued, deviation from equilibrium play cannot be used to directly estimate the ability of players to do iterative reasoning in those games. Auxiliary assumptions such as common knowledge of rationality, common belief in conjectures or common knowledge of the payoffs (possibly included other-regarding concerns), are required. In fact, Weber (2001, footnote 18) reports that some subjects sabotaged others out of spite. It is therefore important to control for social preferences by design. By contrast, in our study, the only cause for subjects deviating from the "equilibrium" (i.e., the optimal sequence of choices in our decision problem) is their own inability to reason.

The studies mentioned above generally conclude that on average individuals behave as if they are able to perform one to two iterations. Given that individuals – due to the nature of the problems discussed – do not necessarily have an incentive to reveal their ability, this conclusion might be too pessimistic.[11]

Another possible approach is to postulates auxiliary assumptions on the behavior and beliefs (types) of individuals, and to estimating the type distributions from data collected in experiments (see, among others, Stahl and Wilson (1994, 1995), Costa-Gomes et al. (2001), Camerer et al. (2004), or Costa-Gomez and Crawford (2006)). For instance, Stahl and Wilson (1994) postulate the following "behavioral" types. The first type $L0$ randomizes uniformly among all his strategies. The second type $L1$ conjectures that his opponent is of type $L0$ and, consequently, best replies to

[8]The same is true for beauty contests where a logically omniscient player chooses the number corresponding to one more iteration step than he believes the others are able to perform. Failing to choose the equilibrium number is not necessarily a sign of limited iterative ability.

[9]Gneezy et al. (2007, 2010) and Dufwenberg et al. (2008) use a version of the game "Nim" to study if and how humans learn backward induction. Since there players have (weakly) dominant strategies, this zero-sum game can be used to infer the depth of counterfactual reasoning from the steps of backward induction performed, if one accepts the auxiliary hypothesis that it is common knowledge that nobody deliberately plays weakly dominated strategies.

[10]In Weber's design, a player correctly inferring his hat's color at stage t was indifferent between revealing his hat color at stage t or at stage $t + 1$.

[11]Considerable cross-game variation also indicates that the inferred ability might not be accurate. See Georganas et al. (2015) for direct evidence.

a uniform distribution over all the actions of his opponents. The third type $L2$ conjectures that his opponent is of type $L1$ and, thus, plays a best-reply to an action of his opponent, which is a best-reply to a uniform distribution over his own action. In general, a player of type Lk plays a best-reply to the best-reply of a type $L(k-1)$.[12] A higher k is thus associated with more sophisticated reasoning. However, the type of an individual is only an imperfect measure of his ability to perform chains of iterative reasoning. Indeed, a type $L2$, for instance, might well be able to perform more sophisticated chains of reasoning, but his conjecture about the play of his opponent implies that he has to perform only a few iterations of reasoning. Observing a type – say $L2$ – just means that the player can at least perform the few iterations needed for $L2$, but may be able to perform many more.[13]

These studies using k-level hierarchies of decision making are very valuable as they provide us with insights into how humans play in games where they need to have models of the rationality and omniscience of others. Recent studies of this kind further improve our understanding of how cognition in specific games takes place by adding more types and either tracking how individuals gather information (Costa-Gomes et al. (2001)) or eliciting beliefs of individuals about the actions of others (Costa-Gomes and Weizsäcker (2008)) in matrix games.[14] However, estimating k-level models involves joint hypotheses and, therefore, cannot directly provide information on the iterative abilities of an individual. Kneeland (2015) shows that in some games it is possible to relax some of the assumptions of level-k models and estimates comparatively high levels of cognitive abilities.

This paper follows yet another route. Our primary question is the following: How many iterations can humans actually do if we remove all strategic uncertainty and control for social preferences? The purpose of this approach is two-fold. Firstly, we want to clarify if humans are actually as limited in their cognitive abilities as previous studies suggest. Secondly, we want to provide the tools and results that can be used to augment the analysis of cognitive decision making in games by providing a measurement of actual levels of "bounded rationality." We believe that this is helpful in order to tackle the question of how much deviation from equilibrium behavior can be attributed to strategic uncertainty or social preferences and how much to the lack of logical omniscience.

There is a growing literature on cognitive abilities and behavior in strategic-form games. For instance, Branãs-Garzaa et al. (2012) show that individuals with higher cognitive abilities as

[12] To control for other-regarding preferences, an "altruistic" type, maximizing the sum of payoffs, is often assumed. However, a wide range of other-regarding preferences (e.g., Fehr and Schmidt (1999) or Charness and Rabin (2002)) are not accounted for.

[13] Camerer et al (2004) consider more sophisticated conjectures: a player of type of Lk conjectures that his opponent might be of any type Lk' with $k' < k$ with strictly positive probability.

[14] Costa-Gomes and Weizsäcker's results are somewhat unsettling, as they find that subjects' actions are not consistent with their (stated) beliefs about the actions of others.

measured by a cognitive reflection test are more likely to play dominant strategies in beauty contest games (the Raven test has no explanatory power, though). Gil and Prowse (2012) show that better cognitive ability, as measured by the Raven test, does not only lead to play closer to equilibrium in beauty contest games, but also to faster learning and convergence to equilibrium if the beauty contests are repeated. See also Branãs-Garzaa et al. (2011), Carpenter et al. (2013), and Putterman et al. (2011) for similar findings on cognitive abilities and behavior in games.

A common feature of this literature is the use of IQ tests or cognitive reflection tests to measure the cognitive abilities of individuals. We believe that our experiment complements these measures in the sense that it provides another measure of the ability of individuals to perform iterative reasoning. A possible advantage of such a measure is that it closely parallels the logic underlying the iterated deletion of dominated strategies. In a companion paper, Bayer and Renou (2016), we construct such a measure, which along with measures of other-regarding preferences, is used to explain the play of individuals in strategic-form games.

The remainder of the paper is organized as follows. Section 2 explains the Red Hat puzzle in detail and highlights the difficulties in implementing it in the laboratory. Section 3 describes our design. Section 4 is devoted to our main results. Section 5 concludes.

2 A simple experiment: the Red Hat puzzle

This section presents a simple puzzle, which will be the basis for our experiment. We follow the exposition of Fagin et al. (1995). Consider N individuals "playing" together. Each of these individuals has either a red hat or a white hat, observes the hat color of others, but cannot observe the color of his own hat. Suppose that some of the individuals, say $n > 0$, have a red hat. Along comes a referee, who declares that "at least one player has a red hat on the head." The referee then asks the following question: "What is your hat color?" All players then simultaneously choose an answer out of "I can't possibly know", "I have a red hat", or "I have a white hat." Players then learn the answers of the other players and are asked again what their hat color is. This process is repeated until all players have inferred their hat color. This problem is known as the Red Hat Puzzle.[15]

We can prove that an individual needs $m + 1$ iteration to figure out his hat color, where m denotes the number of red hats this individual sees. (See Fagin et al. (1995) for a proof.) It is important to stress, however, that the logic for a player correctly inferring his hat color does not rely on the assumption of logical omniscience alone. The remainder of this section examines the

[15]The same game is also known as the "Dirty Faces Game." For an alternative exposition see Fudenberg and Tirole (1991) or Osborne and Rubinstein (1994).

additional assumptions necessary and discusses the difficulties arising with respect to experimental implementation.

Firstly, the logic rests on the assumption of common knowledge of logical omniscience. Even if an individual is logically omniscient, he also needs to know that the answers of the other individuals are logically correct. To see this, suppose that there is a unique red hat and individual 1 observes this red hat. Individual 1 can only correctly infer his hat color (white) if the player, who wears the red hat, answers the first question accurately with "I have a red hat." It follows that any experimental design using the red-hat puzzle to measure the human ability of performing iterated chains of reasoning has to make sure that each individual knows that the answers of other players are logically correct. Otherwise, it would not be possible to separate the effects of subjects' cognitive limitations from those caused by their beliefs about the cognitive abilities of others.

Secondly, the event "There is at least one red hat" must also be common knowledge, otherwise individuals are not able to infer their hat color. To see this, suppose that there is only one red hat. The individual with the red hat observes three white hats, but clearly cannot infer the color of his hat if he does not know that there is at least one red hat. Moreover, even if he knows that there is at least one red hat, the individuals with the white hats cannot infer their hat color if they do not know that the individual with the red hat knows that there is at least one red hat, etc. Our experimental design has therefore to ensure that the event "There is at least one red hat" is common knowledge.

Thirdly, individuals must have incentives to correctly infer their hat color and to truthfully report their logical inferences (even at intermediate steps). Moreover, since an individual's answer influences the subsequent answers of others and, therefore, their payoffs, an individual might want to manipulate his answers to affect the payoff of others if he has other-regarding preferences. Our experimental design will need to properly incentivize individuals and to exclude any confounding influences from social preferences.

3 Experimental protocol and treatments

This section describes our experimental protocol, and how it addresses the difficulties discussed in the previous section.

In our experiment, a human subject was paired with three computers, which were acting as "players." Pairing an individual with computers has several advantages given our objective. Firstly, we can reasonably assume that individuals have no concerns for the eventual "payoffs" of computers. Secondly, we can ensure that a subject knows that the computers' answers are logically correct by a) programming the computer-players to choose the logically correct answers and b) communicating this credibly to the subject. Accordingly, computers were programmed to choose the logically correct answers at each round of questions, and the instructions emphasized this point heavily. The relevant part in the instructions reads: "Recall that the computer-players face the same problem as you do. They can see the hats of all the others but not their own. Therefore, in the above situation [instructions contain a screenshot], Computer 1 knows that the hats of computers' 2 and 3 are white, and also knows your hat colour. However, it does not know its own hat colour. Consequently, the computers also have a logically correct answer to the question: what can you (Computer) infer about your hat colour? The computers ALWAYS choose the logically CORRECT answer." Additionally, subjects were told (and constantly reminded with an on-screen message) that there was at least one red hat.

Note that in game-theoretic terms the explanation given to the subjects does not fully characterize the behavior of the computers. In order to make computers follow the logical solution path it is required that computers have common knowledge of rationality. We decided not to try to explain this in detail to subjects. There are two reasons for it. Firstly, we feared that explaining this could confuse some subjects. Secondly, explaining common knowledge of rationality, contains information that might teach some other subjects how to reason counter-factually. For this reason we chose the term "logically correct" to describe the decision rule. Further below we will show that it is unlikely that subjects misinterpreted how the computers acted and that this is driving results.

Subjects were asked to infer their hat color from the information given to them. At any point when they were asked, they had three possible answers to choose from: "I have a WHITE hat with certainty," "I have a RED hat with certainty," and "I cannot possibly know." The first time a subject had to choose an answer within a puzzle the information a subject had was the hat color of the three computer-players (along with the fact that there was at least one red hat). In any subsequent round within the puzzle, the information a subject had was the complete history of all answers of all players (the computers' and his) in all previous rounds. Similarly, the initial information a computer-player had was the hat color of the two other computers and the human-subject and, subsequently, the complete history of answers. The computers' answers at each point where they were asked was the (unique) logically correct answer inferred from their information and history (assuming that the human player was logically omniscient). Before subjects started the experiment, they had to answer some control questions testing their understanding of the instructions and screen layout (see Figure 3 in the Appendix).

A RHP was stopped after either a wrong answer by the human or a correct announcement of the hat color.[16] This stopping procedure is necessary to avoid logical inconsistencies. Suppose

[16] At a given round, announcing a hat color was correct only if it was actually possible to infer the hat color at this given round.

there is only one red hat, which is worn by the human subject. The subject initially observes three white hats. Now, if the subject (wrongly) answers "I cannot possibly know," then computers should logically infer and, if allowed, answer "I have a red hat." However, this contradicts what the subject observes. Although the computers in this case would choose the logically correct answer, we would have lost control over how a subject interprets this inconsistency. We believe that the observation of contradicting computer announcements and physical reality would have led subjects to believe that the computers were not properly programmed or that our claim that the computers are logically omniscient was based on deception. Stopping a RHP puzzle early does not have any negative consequences for our ability to measure the performance of subjects, as observing the stage where a subject makes the first mistake in a puzzle contains all the information needed.

Since each individual was paired with three computers, we had seven possible distinct logical situations. A logical situation was determined by the number of red hats a subject saw and whether the subject had a red or white hat herself. The more red hats a subject was observing, the more steps (iterations) were required to correctly infer the hat color. We took full advantage of these seven situations to measure the degree of logical omniscience and rationality of individuals.

Treatments. Our experiment consists of six treatments. The treatments differ by the number and order of puzzles presented to the subject, as well as by the feedback given. Puzzles could be ordered in *increasing order of complexity* (i.e. the number of red hats a subject observed was weakly increasing from one puzzle to the next), or the ordering of the puzzles could be random. In some treatments subjects got feedback on whether they solved the previous puzzle correctly or not. In others they did not receive any feedback about previous success or failure. The idea behind varying feedback and order was two-fold. Firstly, this allows us to conduct robustness checks on our measure of logical omniscience. Secondly, we can gain an insight into the conditions under which subjects are able to learn to perform more complicated chains of reasoning.

In treatments I-IV, subjects were asked to play two series of the seven situations presented above. In treatment I, the seven situations were ordered in increasing order of complexity and feedback was provided after each round. In treatment II, the seven situations were also ordered in increasing order of complexity, but no feedback was provided. In treatment III, the seven situations were randomly ordered and feedback was provided. In treatment IV, the seven situations were randomly ordered and no feedback was provided. In treatment V (the one-shot treatment), subjects were asked to play one and only one situation, chosen at random among the seven possible situations. The random draws were independent across subjects. The objective of this additional treatment was to control for pattern recognition and learning within a sequence of the seven situations. Lastly, in treatment VI, subjects were asked to play the seven situations in increasing order of complexity (without feedback) and also several strategic-form games.[17] For convenience, Table 1 summarizes our treatments along with the number of subjects who participated in each treatment.

Treatment	N	**Sequence 1**	**Sequence 2**
I	23	7 RHP ordered, feedback	7 RHP ordered, feedback
II	22	7 RHP ordered, no feedback	7 RHP ordered, no feedback
III	20	7 RHP random, feedback	7 RHP random, feedback
IV	25	7 RHP random, no feedback	7 RHP random, no feedback
V	129	1 (random) RHP	–
VI	30	7 RHP ordered, no feedback	unrelated games

Table 1: Treatments

Payments. Inferring the color of one's hat requires substantial (cognitive) effort. Incentives have therefore to be "powerful" enough for individuals to exert the necessary effort. To provide such powerful incentives, we followed Kahneman and Tversky's (1979) idea that individuals are more sensitive to losses than gains. In all treatments where one or two complete sets of the seven situations were played, subjects started with a lump sum of AU$ 35 (AU$ 17.50 if they were playing only seven situations), and lost AU$ 2.50 for each wrong answer. In treatment V, we used a lottery system, where five winners of substantial prices (AU$ 300 each) were drawn from the pool of subjects who correctly solved their puzzle. The average payment was AU$ 24.89 for treatment I, AU$ 19.54 for treatment II, AU$ 21.12 for treatment III, AU$ 23.3 for treatment IV, AU$ 11.62 for treatment V, and AU$ 8.33 for treatment VI.[18]

The experiments took place at AdLab, the Adelaide Laboratory for Experimental Economics at the University of Adelaide in Australia. We used Urs Fischbacher's (2007) experimental software Z-tree. The 249 participants were mostly students from the University of Adelaide and the University of South Australia.

[17] This treatment was initially used as a pilot for the companion paper. For the present paper, we have pooled the data on the seven situations of the RHP subjects played in treatment VI with the data of treatments I-IV. We conducted our econometric analysis with and without treatment VI and did not find substantial differences. The data from Treatment VI add to the statistical power of our tests.

[18] In treatment VI, subjects were also paid a show-up fee and the payoff received in one of the strategic-form games chosen at random.

4 Empirical results

In this section, we first present empirical regularities regarding the likelihood of subjects correctly inferring their hat color across treatments. We show that the empirical frequency of correctly solved puzzles is relatively high in all treatments, and certainly much higher than we expected.[19] We will also show that in treatments I-IV, the empirical frequency is much higher in the second series of seven puzzles than in the first series, and much higher than in treatment V. This strongly suggests that individuals did learn from introspection and observation. Furthermore, we will see that the likelihood of solving a puzzle requiring three and four iterations are about the same, which suggests that individuals who are able to solve a puzzle requiring three iterations can also solve puzzles requiring four iterations. Later, we report the results of our econometric analysis, which confirms the above empirical regularities and uncovers some other interesting regularities.

4.1 Data analysis

We begin by assessing if there is evidence for subjects' misunderstanding how computers behaved. Recall that we decided not to explain in detail that computers' choices were based on the common knowledge of rationality. So suppose a subject is not sure if a computer chooses based on the assumption of common knowledge of rationality. Then in all puzzles, where a subject herself has a white hat, the computers with red hats will say so before the subject has to switch from "I can't possibly know" to "I have a white hat." Then a person who has worked out the counterfactual reasoning logic can observe that the computers must have assumed rationality of the human players. In the case of a puzzle, where the human has a red hat no such conclusion is possible, since the computers will keep saying "I cannot possibly know" until the human player has worked out her hat color. If there are subjects that are able to iterate but are in doubt about the computers' beliefs, then problems, where subjects have a red hat, should be solved less often than problems with the same number of steps required, where humans have white hats.

In the latter case, a subject who assumed that the computer does not assume common knowledge of rationality will wrongly choose "I cannot possibly know." In the puzzles, where the human has a white hat, there is proof that the computer has chosen based on the assumption of common knowledge of rationality and therefore should not make such a mistake. Looking at the one-shot treatment where learning or order effects don't play a role, shows that our subjects do not make more mistakes in the puzzles where they don't have an indication that computers choose under the assumption of common knowledge of rationality.[20]

[19] We actually had to run to the bank during the first sessions to get more cash!

[20] The one-sided test that the proportion of correct answers is greater with white hats is rejected with p-values of

We first report the percentage of individuals over the entire sample who are still correct after m' ($m' = 0, \ldots, m+1$) iterations (steps) of reasoning when $(m+1)$ iterations are needed to correctly infer one's hat color (see Figure 1). For instance, in the second panel of Figure 1, individuals needed to perform two iterations of reasoning to correctly infer their hat color. About 91.09% of the answers were still correct at the first iteration (i.e., 91.09% of the individuals correctly performed one step out of the two steps needed), and 75.43% were correct overall (i.e., performed two steps). Furthermore, in each panel, the bar furthest to the right represents the percentage of individuals who correctly solved a puzzle (as a function of the number of steps needed to solve it). If one step was needed, 99.13% of the answers were correct, 75.43% if two steps were needed, 44.05% if three steps were needed and 39.87% if four steps were needed.[21]

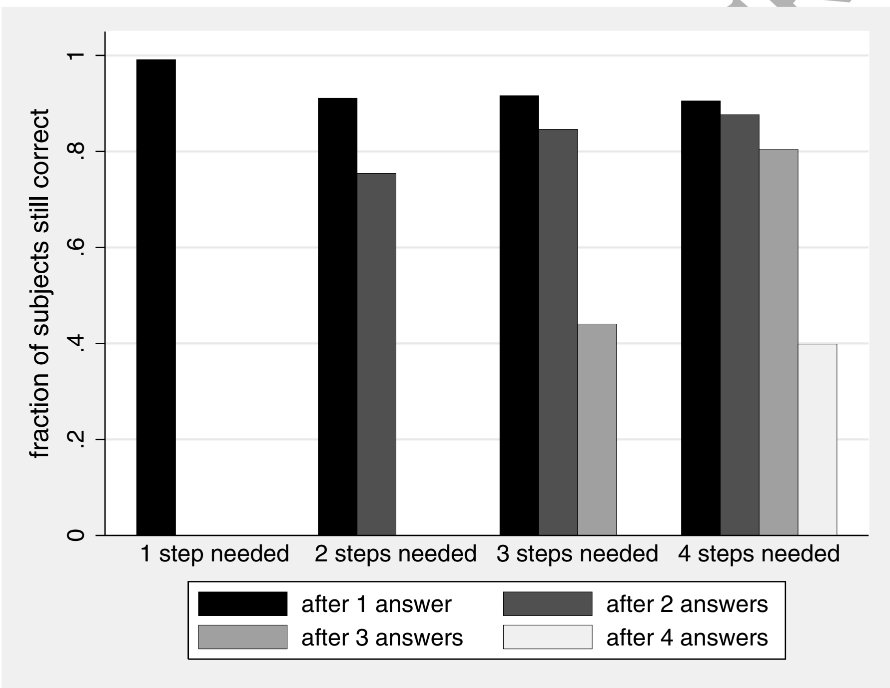


Figure 1: Frequency of correct answers

Several observations are worth making. Firstly, as expected, the more iterations are required to correctly infer one's hat color, the lower is the percentage of correct answers. Secondly, and somewhat surprisingly, there is almost no difference between solving a puzzle requiring three iterations and one requiring four iterations (44.05% vs. 39.87%), while there is a sizeable difference between

$p > 0.17$ for two, $p > 0.88$ for three and $p > 0.81$ for four steps.

[21]Conditioning on the individuals who correctly answered our control questions, the percentages become 99.45 %, 78.59%, 46.39% and 44.23%, respectively.

solving a puzzle requiring two iterations and one requiring three or four. The pairwise correlations reported in Table 2 reinforce this observation. The within-subject correlation between the number of correct answers when two and three steps are required is about the same as the within-subject correlation between the number of correct answers when two and four steps are required (0.423 vs. 0.425), while the correlation between the number of correct answers when three and four steps are required is significantly higher at 0.641.[22] This implies that a subject, who can solve the two-step puzzles, is equally likely to solve puzzles with three and four steps. Moreover, the correlation between solving three-step and four-step puzzles is highest. A large number of subjects are able to solve both types of puzzles.

	# correct 1 step	# correct 2 steps	# correct 3 steps	# correct 4 steps
# correct 1 step	1.000			
# correct 2 steps	0.058	1.000		
# correct 3 steps	0.048	0.423***	1.000	
# correct 4 steps	−0.016	0.425***	0.641***	1.000

Table 2: Correlation table for the number of puzzles correctly solved.

Thirdly, the largest drop in the empirical distributions in Figure 1 is occurring at the *critical iteration*, which occurs at the $m+1$st decision. Note that up to the mth announcement the correct answer is "I can't possibly know." Many subjects might have chosen this correct answer for the wrong reason. A subject might simply fail to grasp the logical implications of his knowledge and, consequently, choose to answer "I can't possibly know" because of his confusion. Intermediate steps might therefore reflect very biased information about the logical omniscience of individuals. For this reason, we will ignore intermediate steps for our econometric analysis.

Figure 2 presents the percentage of puzzles correctly solved as a function of the number of iterations required and whether individuals have played them in isolation (treatment V), for the first (1st seq.) or second time (2nd seq.) in the treatments with repetition (treatments I-IV). A striking pattern emerges: individuals do seem to learn from both introspection and observation. If we compare individuals playing a puzzle of a certain complexity in isolation (treatment V) with individuals playing the same situation for the first time in treatments I-IV, the likelihood of solving

[22]Testing for the equality of the correlation coefficients (using a test based on the Fisher z-transformation) between two and three steps and two and four steps, we do not reject the null hypothesis, while we do reject the null hypothesis that correlation coefficients between two and three steps (or two and four steps) and three and four steps are the same.

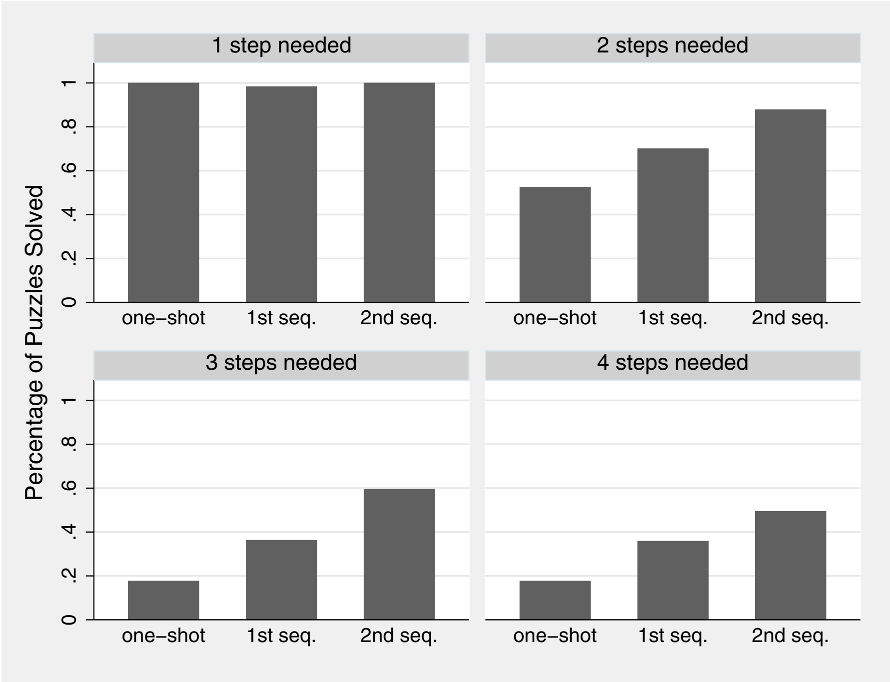


Figure 2: Correct answers by steps necessary and repetition

the puzzle is much higher in the latter treatments. Moreover, the likelihood of individuals correctly solving the same puzzle the second time round is even higher.[23] This is true in all treatments with repetition. It seems somewhat surprising that the same pattern emerges if we look at treatments II and IV, where individuals did not receive feedback. In that case, learning seems to mainly come from introspection (since no hard information is provided). Note, however, that in treatments II and IV with no feedback, some inference was nonetheless possible when the game was stopping after the announcement "I cannot possibly know." An individual might have inferred that he should have known his hat color. However, a subject can only rely on her own understanding of the problem to make an inference in such a situation. Introspection is a potential reason for improved performance while experimentation is another. In contrast, in treatments I and III, where feedback about having solved the puzzle correctly or not was provided, subjects did not have to infer if they made a mistake. This makes learning easier but also gives a cue for subjects that made mistakes to experiment with alternative strategies.

In treatments II and IV, there are 9 and 10 observations, respectively, of an individual incorrectly

[23]Over the entire sample, there are only four observations of an individual correctly solving a puzzle for the first time and failing the second time. Surprisingly, the four observations are in treatments I and III, the treatments with feedback.

answering "I cannot possibly know" the first time he plays a puzzle and correctly solving the puzzle the second time. Over-all, the number of observations of an individual incorrectly answering "I cannot possibly know" the first time a puzzle was played is 21 and 19, respectively. In comparison, in treatments I and III with feedback, the respective numbers of observations are 6 out of 12 and 6 out of 10. So the likelihood to correctly infer one's hat color the second time round conditional on incorrectly answering "I cannot possibly know" the first time is 0.5 in treatment I, 0.43 in treatment II, 0.6 in treatment III and 0.53 in treatment IV. This indicates that subjects made similar inferences in treatments with feedback and without feedback when the game was stopping after the announcement "I cannot possibly know."

We can contrast these numbers with the numbers of observations of individuals incorrectly stating their hat color the first time they play a puzzle and correctly solving the puzzle the second time. In treatments II and IV with no feedback, the numbers are 8 out of 30 and 2 out of 19, respectively.[24] In comparison, in treatments I and III with feedback, the respective numbers of observations are 9 out of 20 and 5 out of 22. We can note that the likelihood to correctly infer one's hat color the second time round conditional on incorrectly stating it the first time is about twice as high in treatment I than in treatment II (0.45 vs 0.26) and also in treatment III than in treatment IV (0.22 vs 0.10). This implies that the improvement of performance due to repetition without feedback is almost entirely driven by the cases, where subjects ended without stating a hat color the first time around. These are cases where there was the possibility to infer that something must have been wrong.

4.2 Econometric analysis

We now consider the determinants of correctly inferring one's hat color more formally. The dependent variable in our regressions is "correct," a dichotomous variable indicating whether a subject had correctly inferred his hat color in a given puzzle. We estimated four different econometric models: a probit and a logit model allowing for error clustering within subjects, and a probit and logit model allowing for random intercepts for individuals (i.e., panel models with subject-specific random effects). We report these four regressions to assess the robustness of our analysis. We ran logit and probit models to test whether our results are robust to the specifications of the response probability. The naturally preferred econometric models are the panel models (we have observations for the same individual over time). However, the panel specification might suffer from biases due to the unbalanced panel structure. We either observe an individual's answer(s) in one puzzle

[24]The second number is the number of observations of individuals incorrectly stating their hat color the first time they play a puzzle.

(treatment V), seven puzzles (treatment VI) or fourteen puzzles (treatments I-IV).[25] We therefore report the results from cross-sectional regressions with error clustering (pooled probit and logit) as a robustness check.

We used treatment I (ordered, feedback) as the reference group for the treatment dummies, since it has the highest success rates. Table 3 reports the marginal effects averaged over the whole sample.

Table 3: Determinants of correct choices in the Red-Hat Puzzle

	Probit	Logit	Panel Probit	Panel Logit
Avg. marg effects	**Dependent variable:** *correct* (situation correctly solved)			
Iteration step dummies (1 step is the reference in all regressions)				
2 steps needed	-0.218^{***}	-0.221^{***}	-0.184^{***}	-0.179^{***}
	(0.024)	(0.024)	(0.025)	(0.026)
3 steps needed	-0.540^{***}	-0.545^{***}	-0.558^{***}	-0.563^{***}
	(0.030)	(0.029)	(0.036)	(0.038)
4 steps needed	-0.603^{***}	-0.601^{***}	-0.626^{***}	-0.631^{***}
	(0.034)	(0.033)	(0.035)	(0.037)
Treatment dummies (ordered, feedback is the reference)				
ordered, no feedback	-0.180^{***}	-0.181^{***}	-0.199^{***}	-0.199^{***}
	(0.055)	(0.055)	(0.067)	(0.070)
random, feedback	-0.081	-0.083	-0.101	-0.106
	(0.063)	(0.063)	(0.080)	(0.084)
random, no feedback	-0.095	-0.094	-0.129^{*}	-0.135^{*}
	(0.072)	(0.071)	(0.077)	(0.081)
one shot	-0.251^{***}	-0.257^{***}	-0.291^{***}	-0.301^{***}
	(0.058)	(0.059)	(0.068)	(0.071)
Degree dummies (economics is the reference group)				
engineering	0.170^{**}	0.157^{*}	0.118	0.115
	(0.085)	(0.085)	(0.109)	(0.113)
medicine	0.238^{**}	0.225^{**}	0.189^{*}	0.184
	(0.099)	(0.099)	(0.127)	(0.132)
Other degrees (arts, commerce, finance, science, law) not significant				

[25]We did the econometric analysis with treatments I-IV only and did not find qualitative differences.

Table 3: ... continued

	Probit	Logit	Panel Probit	Panel Logit
gender (male=1)	0.139***	0.142***	0.171***	0.179***
	(0.041)	(0.041)	(0.049)	(0.051)
control questions OK	0.079	0.078	0.130**	0.141**
	(0.056)	(0.057)	(0.056)	(0.059)
repetition	0.096***	0.091***	0.152***	0.148***
	(0.023)	(0.023)	(0.024)	(0.025)
time first choice	−0.003***	−0.004***	−0.0001	−0.0001
	(0.001)	(0.001)	(0.001)	(0.001)
Dummies for age, advanced math, critical time, all not significant				
N	1599	1599	1599	1599
ρ	–	–	0.504***	0.517***
Log-likelihood	−745.15	−743.55	−636.97	−630.79

Standard errors in parentheses; *** p <0.01, ** p<0.05, * p<0.1; $Prob>\chi^2<0.0001$ (all models)

We organize our econometric analysis around two main themes: iterative reasoning and learning.

Iterative reasoning. All our regressions confirm our initial observation: the more steps required to solve a puzzle, the less likely it is that an individual solves it ($p < 0.01$, Wald tests in all applicable models). For instance, with the Panel Probit model, the probability of solving a puzzle requiring two steps is 0.184 lower than that for the puzzle requiring one step (the reference). The probability difference to the one-step puzzle is even lower if three or four steps are required (lower by 0.558 and 0.626, respectively). The predicted probability of correctly inferring one's hat color when two steps are required are about twice as high as that for solving a puzzle requiring four steps (81 percent vs. 37 percent according to the probit panel model). All four econometric models produce similar predictions. However, the likelihood does not differ much between three and four steps (e.g. 37 versus 44 percent in the Panel Probit model). Subjects who understand how to solve the puzzle with three steps are likely to have understood the general principle and therefore are also able to solve puzzles of difficulty four. While the different predicted probabilities only show small differences, these differences are significant though.[26] We reject the null hypothesis that solving a puzzle requiring two steps is equivalent to one requiring three or four steps in all models with even more confidence ($p < 0.01$, Wald tests).

Two closely related experimental studies are Weber (2001) and Bayer and Chan (2009). Weber

[26]Wald test in all models. The p-values range from 0.032 to 0.037 depending on the model.

implements the red hat puzzle as a dynamic game of incomplete information between two or three human players. Bayer and Chan replicate Weber's experiment and compare the replication with a modified version of Weber's game, where equilibria in weakly dominated strategies are eliminated.[27] As already argued, deviation from equilibrium play cannot be used to directly estimate the ability of players to do iterative reasoning in those games. Auxiliary assumptions such as common knowledge of rationality, common belief in conjectures or common knowledge of the payoffs (possibly other-regarding preferences), are required. By contrast, in our study, the only cause for subjects deviating from the "equilibrium" (i.e., the optimal sequence of choices in our decision problem) are their own inability to reason.

Consequently, comparisons between these studies and our study are difficult. The most severe problem is related to the fact that in the interactive version of the red hat puzzle, a premature end of the game is possible. All players can end the game at each stage by playing the action "down," which is equivalent to choosing the statement "I have a red hat" in our setting. Whenever a player erroneously (i.e., deviating from the equilibrium path) chooses this action, the game ends prematurely. Therefore, it is impossible to assess whether the other players would have played according to the equilibrium strategies in later stages. There are two reasonable ways of dealing with this problem. Weber implicitly assumes that players, who had followed the equilibrium path up to the stage where the game ended prematurely, would have continued on the equilibrium path in later stages. This overestimates equilibrium agreement. This bias is likely to be large, especially if we consider our findings (see Figure 1) that the last choice a subject has to make is the critical choice and most mistakes are made at this stage.

A second approach (Bayer and Chan) is to drop all individual observations, where other players prematurely ended the game. Results based on this second approach are better suited for comparison with the results of our study, as in our experiment the other players (the computers) never prematurely end the game. Using this approach, Bayer and Chan report agreement rates with sequential equilibrium for their three player red-hat game of 86.7 percent when one step of iteration is needed, 10.6 percent when two steps are needed, and 5.3 percent when three steps are needed. In a modified version of the game, which ensures uniqueness of the equilibrium, the agreement is higher (91.7 percent when one step is needed, 39.7 percent when two are and 19.1 percent when three steps are needed). The success rates in our experiment with 99.7, 81.3 and 43.9 percent (predicted by the panel probit for one, two and three steps) are clearly higher than in the studies mentioned above.[28] This suggests that a high proportion of the deviations from "equilibrium behavior" observed in

[27] In Weber's design, a player correctly inferring his hat's color at stage t was indifferent between revealing his hat color at stage t or at stage $t+1$.

[28] Note that the standard errors of our predictions is quite low and range from 0.002 (one step) to only 0.036 (three steps), which indicates that the differences are not likely to stem from chance.

the studies of Weber (2001) and Bayer and Chan (2009) might have resulted from players' doubts about the rationality of others and not from their own limited ability.

Beside the impact of the level of difficulty, we found a strong gender effect. Male subjects did significantly better. The average likelihood of solving a puzzle for a male was – depending on the model used – between 0.139 and 0.178 higher than that for a female. Quite intuitively, once one allows for unobserved heterogeneity, having made a mistake in the control questions reduces the likelihood of solving a puzzle correctly by between 0.13 and 0.14.

Learning. The results reported in Table 3 provide some insights into learning. We find both learning through repetition and learning within a sequence across all treatments. The surprising observation here is that explicit feedback about being correct or not is not necessary for learning. Introspection and the limited feedback of games ending prematurely in situations where some specific errors are made are sufficient.

Holding everything else (i.e. treatment, difficulty of the puzzle and subject characteristics) constant, repetition increases the probability of a success considerably. Depending on the model used, the estimated average increase of the success probability is between 0.091 and 0.152, when the same puzzle is played the second time. An alternative regression (see Table 4 in the online Appendix), where we use interaction dummies between repetition and treatment, shows that this learning effect from playing a puzzle a second time is present regardless of the treatment. The success probabilities go up by between 0.09 (random without feedback) and 0.19 (ordered without feedback) when a puzzle is repeated. The increases are all significantly different from zero on the 5% level according to Wald tests. Consequently, learning in our game does not require feedback or a specific order of the puzzles. In fact, the change in probabilities between the first and the second time a puzzle was played, are not significantly different across treatments (Wald tests).

Further, we find evidence that subjects learn from being exposed to one puzzle how to solve a different puzzle. Recall that in all treatments but the one-shot treatment, subjects played all seven puzzles before repeating. The coefficients for the treatment dummies (in Table 3 measure the difference in the probability of solving a puzzle in the corresponding treatment for given difficulty of a puzzle and and the first time the puzzle was encountered. So differences in the treatment dummies' coefficients between one-shot and other treatments can be used as evidence for learning across puzzles. We find that subjects do better if the puzzle is part of a sequence than if it is played one-shot (significant on the 2.5% level for all treatments, Wald tests based on either the Probit or Logit panel models.)

5 Conclusion

This paper argues that using dominance-solvable games (such a beauty contests, dirty-faces games or centipede games) in order to measure human's iterative abilities overlooks that behaviour in these games is not only the consequence of iterative abilities but is also influenced by beliefs about the rationality of others and by social preferences. We demonstrate this by running a series of novel experiments on the dirty-faces game (here called the Red-Hat Puzzle), where we are able to control for the other influences. We show by comparing our results to those from previous studies using this game (Weber, 2001 and Bayer and Chan, 2009) that subjects are able to perform more steps of counterfactual reasoning than initially thought.

Moreover, we argue that our setup can be used to obtain a measure of iterative reasoning abilities, which has the two advantages that a) is not confounded with anything else and b) is still derived from a game-like situation. Our measure can be used as an explanatory variable in studies that have the aim of separating the effect of cognitive abilities on play in games from the impact of strategic uncertainty and social preferences.

References

[1] Marina Agranov, Elizabeth Potamites, Andrew Schotter and Chloe Tergiman (2012), Beliefs and endogenous cognitive levels: An experimental study, Games and Economic Behavior, 75, 449-463.

[2] Larbi Alaoui and Antonio Penta (2015), Endogenous Depth of Reasoning, forthcoming in Review of Economic Studies

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

[4] Ralph-C Bayer and Mickey Chan (2009), The Dirty Faces Game Revisited, School of Economics Working Paper, The University of Adelaide.

[5] Ralph-C. Bayer and Ludovic Renou (2016), Cognitive abilities and behavior in strategic-form games, forthcoming in Journal of Economic Psychology.

[6] Ralph-C Bayer and Ludovic Renou (2007), Measuring the depth of iteration in humans in L. Oxley and D. Kulasiri (Eds.), Proceedings of the International Congress on Modelling and Simulation 2007, 379-385.

[7] Pablo Branãs-Garzaa, Teresa García-Munõza, and Roberto Hernán Gonzáleza (2012), Cognitive Effort in the Beauty Contest Games. Journal of Economic Behavior and Organization, 83, 254-260.

[8] Pablo Branãs-Garzaa, María Paz Espinosa, and Pedro Rey-Biel (2011), Travelers' Types, Journal of Economic Behavior & Organization 78, 25-36

[9] Randolph T. Beard and Richard Beil (1994), Do People Rely on the Self-interested Maximization of Others? An Experimental Test, Management Science, 40, 252-262.

[10] Adam Brandenburger, Amanda Friedenberg and H. Jerome Keisler (2008), Admissibility in Games, Econometrica, 76, 307-352.

[11] Susana Cabrera, C. Mónica Capra and Rosarie Gómez (2006), Behavior in One-shot Traveler's Dilemma Games: Model and Experiment with Advise, Spanish Economic Review, 9(2), 129-152.

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

[13] Colin F. Camerer (2003), *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton University Press, Princeton, USA.

[14] Jeffrey Carpenter, Michael Graham, Jesse Wolf (2013), Cognitive Ability and Strategic Sophistication, Games and Economic Behavior 80, 115-130.

[15] Gary Charness and Matthew Rabin (2002), Understanding Social Preferences with Simple Tests, Quarterly Journal of Economics 117, 817-869.

[16] Peter Fleming and Daniele Zizzo, A Simple Stress Test of Experimenter Demand Effects, Theory and Decision, forthcoming, 2015.

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

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

[19] Miguel A. Costa-Gomes and Georg Weizsäcker (2008), Stated Beliefs and Play in Normal-Form Games, Review of Economic Studies 75, 729-762.

[20] Martin Dufwenberg, Ramya Sundaram and David Butler (2008), Epiphany in the Game of 21, Journal of Economic Behavior and Organization Vol.75, 132-143

[21] David Gill and Vicoria Prowse (2015), Cognitive Ability, Character Skills, and Learning to Play Equilibrium: A Level-k Analysis, forthcoming in the Journal of Political Economy.

[22] Uri Gneezy, Aldo Rustichini and Alexander Vostroknutov (2007), I'll Cross That Bridge When Come to It: Backward Induction as a Cognitive Process, mimeo.

[23] Gneezy, Uri Gneezy, Aldo Rustichini and Alexander Vostroknutov (2010). Experience and Insight in the Race Game. Journal of Economic Behavior Organization 75, 144-155

[24] Ronald Fagin, Joseph Y. Halpern, Yoram Moses and Moshe Y. Vardi (1995), *Reasoning about Knowledge*, MIT Press, Cambridge, USA.

[25] Ernst Fehr and Klaus Schmidt (1999), A Therory of Fairness, Competition, and Cooperation, Quarterly Journal of Economics 114, 817-868.

[26] Urs Fischbacher (2007), z-Tree: Zurich Toolbox for Ready-made Economic Experiments, Experimental Economics 10(2), 171-178.

[27] Drew Fudenberg and Jean Tirole (1991), *Game Theory*, MIT Press, Cambridge, USA.

[28] Sotiris Georganas, Paul Healy and Roberto Weber (2015), On the Persistence of Strategic Sophistication, Journal of Economic Theory 159, 369-400.

[29] David Gil and Victoria Prowse, (2013), Cognitive Ability and Learning to Play Equilibrium: A Level-k Analysis. Mimeo, Oxford.

[30] Jacob K. Goeree and Charles A. Holt (2001), Ten Little Treasures of Game Theory and Ten Intuitive Contradictions, American Economic Review 91, 1402-1422.

[31] Eric J. Johnson, Colin Camerer, Sankar Sen, Talia Rymon (2002), Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining, Journal of Economic Theory 104, 16-47.

[32] Teck Ho, Colin F. Camerer and Keith Weigelt (1998), Iterated Dominance and Iterated best-response in Experimental p-beauty Contests, American Economic Review, 88, 947-969.

[33] Daniel Kahneman and Amos Tversky, (1979), Prospect Theory : An Analysis of Decision under Risk, Econometrica 47, 263-291.

[34] Terri Kneeland (2015), Identifying High-Order Rationality, forthcoming in Econometrica.

[35] Bart Lipman (1999), Decision without Logical Omniscience: Toward an Axiomatic Framework for Bounded Rationality, Review of Economic Studies 66, 339-361.

[36] Richard D. McKelvey and Thomas R. Palfrey (1992), An Experimental Study of the Centipede Game, Econometrica 58, 1321-1339.

[37] C. Nicholas Mc Kinney and John Van Huyck (2007), Estimating Bounded Rationality and Pricing Performance Uncertainty, Journal of Economic Organization and Behavior 62, 625-639.

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

[39] Martin J. Osborne and Ariel Rubinstein (1994), *A Course in Game Theory*, MIT Press, Cambridge, USA.

[40] Louis Putterman, Jean-Robert Tyran and Kenju Kamei (2011), Public Goods and Voting on Formal Sanction Schemes, Journal of Public Economics 95, 1213-1222.

[41] Dale O. Stahl and Paul W. Wilson (1994), Experimental Evidence on Players' Models of Other Players, Journal of Economic Behavior & Organization 25, 309-327.

[42] Dale O. Stahl and Paul W. Wilson (1995), On Players' Models of Other Players: Theory and Experimental Evidence, Games and Economic Behavior 10, 218-254.

[43] Alan Turing (1950), Computing Machinery and Intelligence, Mind 49, 433-460.

[44] John B. Van Huck, John Wildenthal and Raymond C. Battalio (2002), Tacit Cooperation, Strategic Uncertainty, and Coordination Failure: Evidence from Repeated Dominance Solvable Games, Games and Economic Behavior 38, 156-175.

[45] Roberto A. Weber (2001), Behavior and Learning in the Dirty Faces Game, Experimental Economics 4, 229-242.

[46] Roberto A. Weber (2003), 'Learning with No Feedback in a Competitive Guessing Game, Games and Economic Behavior, 44, 134-144.

A Appendix not for publication

This material is planned to go into an online appendix

A.1 Additional regression

Table 4: Learning: observation and introspection

	Panel Probit	Panel Logit
Avg. marg effects	**Dependent variable**: *correct*	
Iteration step dummies (1 step is the reference in all regressions)		
2 steps needed	-0.183^{***}	-0.178^{***}
	(0.025)	(0.025)
3 steps needed	-0.557^{***}	-0.562^{***}
	(0.036)	(0.038)
4 steps needed	-0.625^{***}	-0.631^{***}
	(0.035)	(0.037)
Treatment and sequence dummies (one-shot is the reference)		
ordered, feedback, seq. 1	0.283^{***}	0.293^{***}
	(0.070)	(0.073)
ordered, feedback, seq. 2	0.444^{***}	0.453^{***}
	(0.071)	(0.074)
ordered, no feedback, seq. 1	0.082	0.095
	(0.060)	(0.062)
ordered, no feedback, seq. 2	0.272^{***}	0.274^{***}
	(0.069)	(0.071)
random, feedback, seq. 1	0.172^{**}	0.180^{**}
	(0.078)	(0.081)
random, feedback, seq. 2	0.340^{***}	0.350^{***}
	(0.077)	(0.080)
random, no feedback, seq. 1	0.185^{***}	0.191^{***}
	(0.070)	(0.073)
random, no feedback, seq. 2	0.277^{***}	0.279^{***}
	(0.071)	(0.073)
gender (male=1)	0.170^{***}	0.179^{***}
	(0.049)	(0.051)
control questions OK	0.131^{**}	0.142^{**}
	(0.056)	(0.058)

Table 4: . . . continued

	Panel Probit	Panel Logit
All other dummies for ages, degrees and decision times not significant		
N	1599	1599
ρ	0.505***	0.517***
Log-likelihood	−635.55	−629.37

Standard errors in parentheses; *** p <0.01, ** p<0.05, * p<0.1;

$Prob{<}\chi^2{<}0.0001$ (all models)

A.2 Control questions

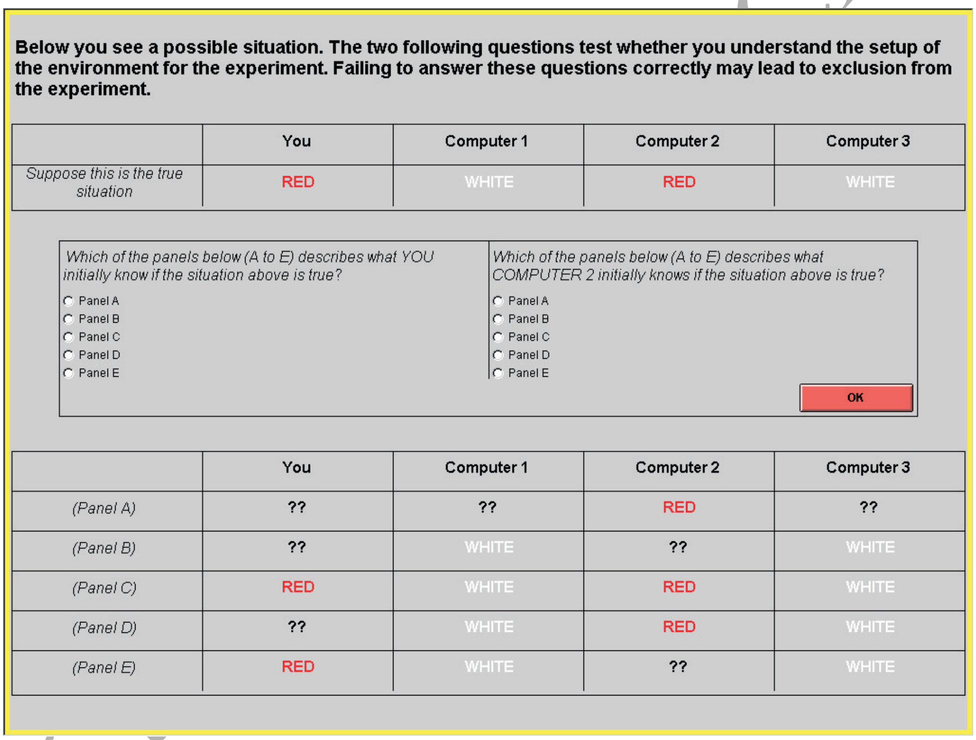


Figure 3: Screenshot of control questions

A.3 Instructions: Treatment V (one-shot)

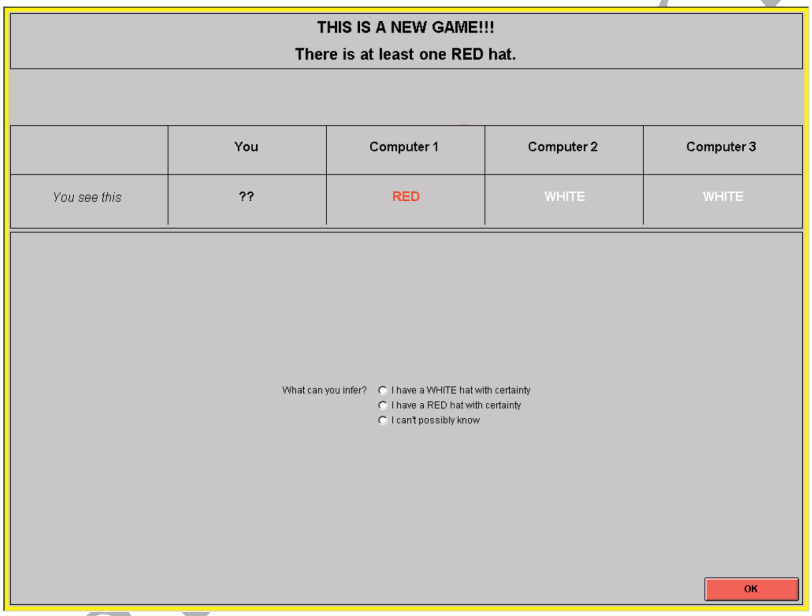
Instructions

Thank you for your participation in this experiment. If you read these instructions carefully and act upon them, you can earn real money.

You are not allowed to communicate with other participants during the course of the experiment. If you do not follow this rule you may be excluded from the experiment.

Your task

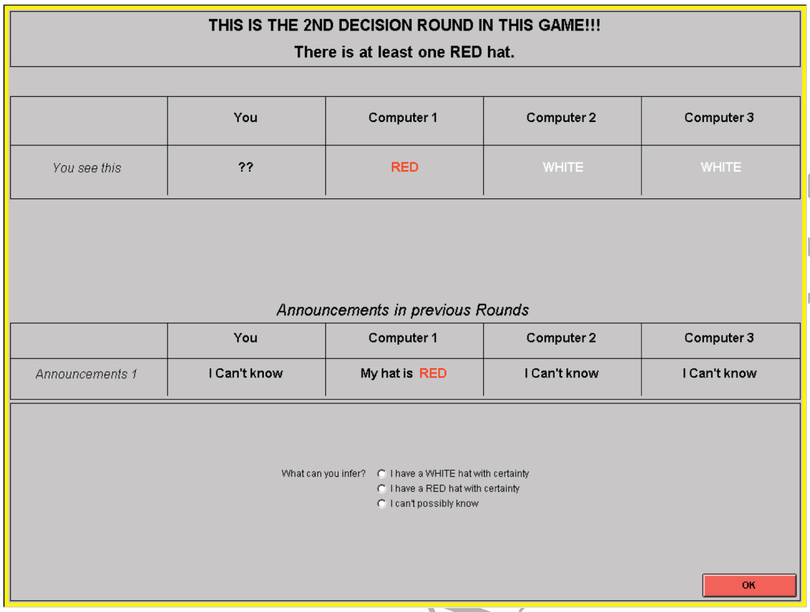
Your task in this experiment is to determine the colour (red or white) of your hat. You will be paired with 3 computer-players. You will be able to see the colour of the hats of the computer-players, but **not** the colour of your own hat. The computer-players are in a similar situation. They observe your hat colour and the hat colours of their fellow computer-players, but not the colour of their own hat. However, everybody knows (you and the computer-players) that *at least one player has a red hat*. The picture below shows a typical situation:



You observe in this case that one of the computer-players has a red hat while the other two have white hats. The question marks "??" indicate that you do not know your hat colour.
You are asked to decide what you can infer from the information you are given. Possible answers are: "I have a WHITE hat with certainty", "I have a RED hat with certainty", and "I can't possibly know". One of these answers is correct, the two others are wrong. Note that answering "I can't possibly know" is wrong whenever it is possible to correctly infer the hat colour from the information given. Similarly ticking "I have a WHITE hat with certainty" or "I have a RED hat with certainty" is only correct if it is actually possible to logically infer that your hat colour is white or red.

The game may end after your initial decision. If the game continues, you will be given the additional information of what the computer-players have inferred from their observation.

Recall that the computer-players face the same problem as you do. They can see the hats of all the others but not their own. Therefore, in the above situation, Computer 1 knows that the hats of computers' 2 and 3 are white, and also knows your hat colour. However, it does not know its own hat colour. Consequently, the computers also have a logically correct answer to the question: what can you (Computer) infer about your hat colour? The computers ALWAYS choose the logically CORRECT answer.



Above you can see a possible screen for your second decision. You again have to decide what you can infer about your hat's colour. However, now you have the additional information about what the computers (correctly) announced in the decision round before. After you have made another decision, the game may end or continue. If the game continues, you will again be given the additional information of what the computers inferred from the previous round. This process will go on until you either correctly inferred your hat colour or until you made a mistake.

Payment
You will play one of these games. If you solve your puzzle and correctly determine your hat colour then you will be put in the draw for a prize of AUD 300. The draw will be conducted later this year, when the whole series of experiments has been conducted.

Introductory questions
Before you start the actual game we will ask you some questions about the game. These questions will be designed to test if you understand the instructions. Please make sure to read the instruction very carefully, as failing to answer the pre-game questions correctly may lead to exclusion from the experiment.

Questions
Do you have any questions? If yes please raise your hand and we will come and answer them in private.

A.4 Instructions: Treatment VI (without repetition)

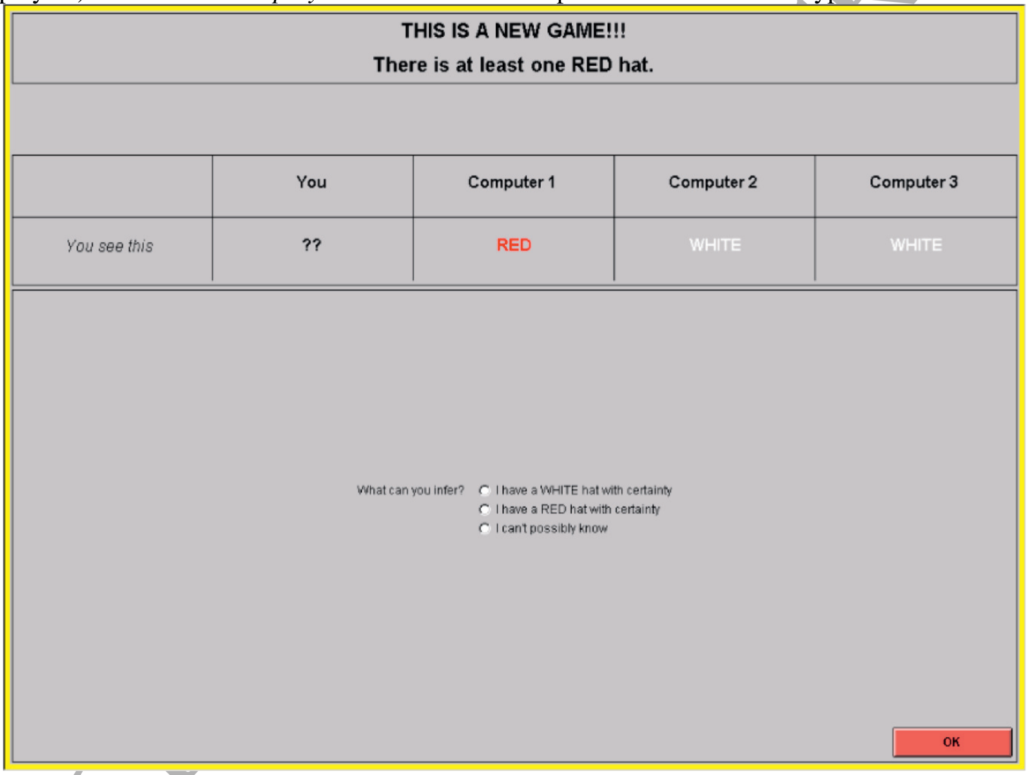
Instructions

Thank you for your participation in this experiment. If you read these instructions carefully and act upon them, you can earn real money.

You are not allowed to communicate with other participants during the course of the experiment. If you do not follow this rule you may be excluded from the experiment.

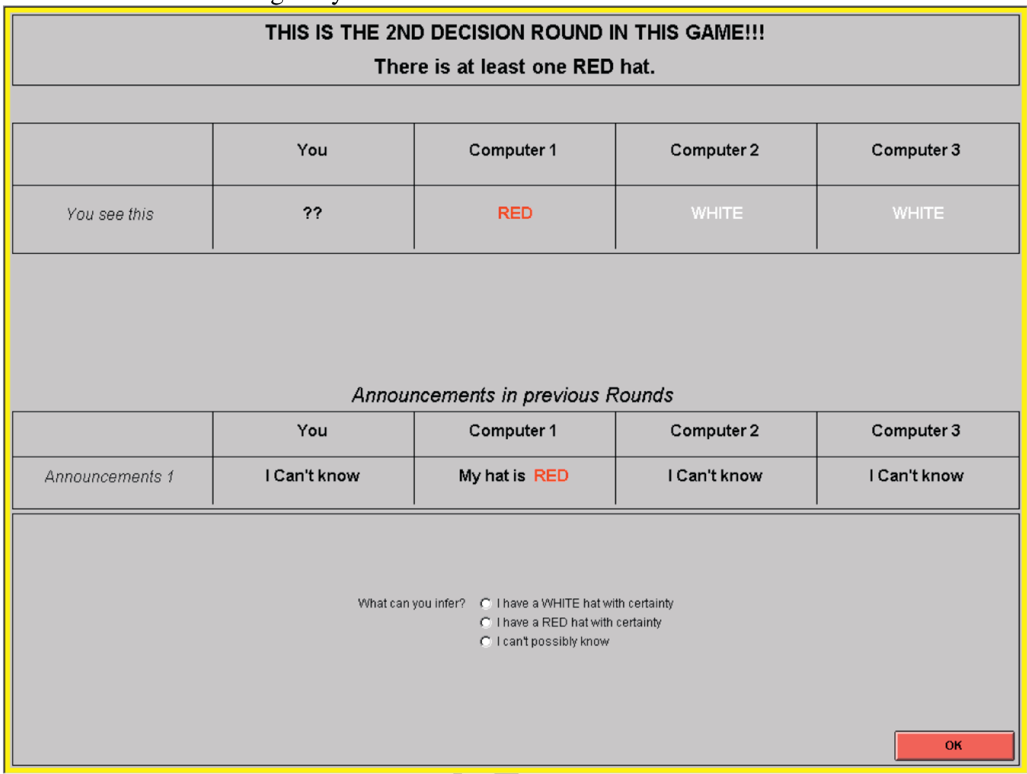
Your task

Your task in this experiment is to determine the colour (red or white) of your hat. You will be paired with 3 computer-players. You will be able to see the colour of the hats of the computer-players, but **not** the colour of your own hat. The computer-players are in a similar situation. They observe your hat colour and the hat colours of their fellow computer-players, but not the colour of their own hat. However, everybody knows (you and the computer-players) that *at least one player has a red hat*. The picture below shows a typical situation:



You observe in this case that one of the computer-players has a red hat while the other two have white hats. The question marks "??" indicate that you do not know your hat colour.
You are asked to decide what you can infer from the information you are given. Possible answers are: "I have a WHITE hat with certainty", "I have a RED hat with certainty", and "I can't possibly know". One of these answers is correct, the two others are wrong. Note that answering "I can't possibly know" is wrong whenever it is possible to correctly infer the hat colour from the information given. Similarly ticking "I have a WHITE hat with certainty" or "I have a RED hat with certainty" is only correct if it is actually possible to logically infer that your hat colour is white or red.

The game may end after your initial decision. If the game continues, you will be given the additional information of what the computer-players have inferred from their observation. Recall that the computer-players face the same problem as you do. They can see the hats of all the others but not their own. Therefore, in the above situation, Computer 1 knows that the hats of computers' 2 and 3 are white, and also knows your hat colour. However, it does not know its own hat colour. Consequently, the computers also have a logically correct answer to the question: what can you (Computer) infer about your hat colour? The computers ALWAYS choose the logically CORRECT answer.



Above you can see a possible screen for your second decision. You again have to decide what you can infer about your hat's colour. However, now you have the additional information about what the computers (correctly) announced in the decision round before. After you have made another decision, the game may end or continue. If the game continues, you will again be given the additional information of what the computers inferred from the previous round. This process will go on until you either correctly inferred your hat colour or until you made a mistake.

<u>*Different games*</u>
After a game has ended, another new game will start. You will be given a set of 7 games.

<u>*Payment*</u>
You will start with 20 Australian Dollars. For each mistake you make, we will deduct 2.50 Dollars from your account. After the 7 games you will be paid the amount remaining in cash.

<u>*Introductory questions*</u>
Before you start the actual game we will ask you some questions about the game. These questions will be designed to test if you understand the instructions. Please make sure to read the instruction very carefully, as failing to answer the pre-game questions correctly may lead to exclusion from the experiment.

<u>*Questions*</u>
Do you have any questions? If yes please raise your hand and we will come and answer them in private.

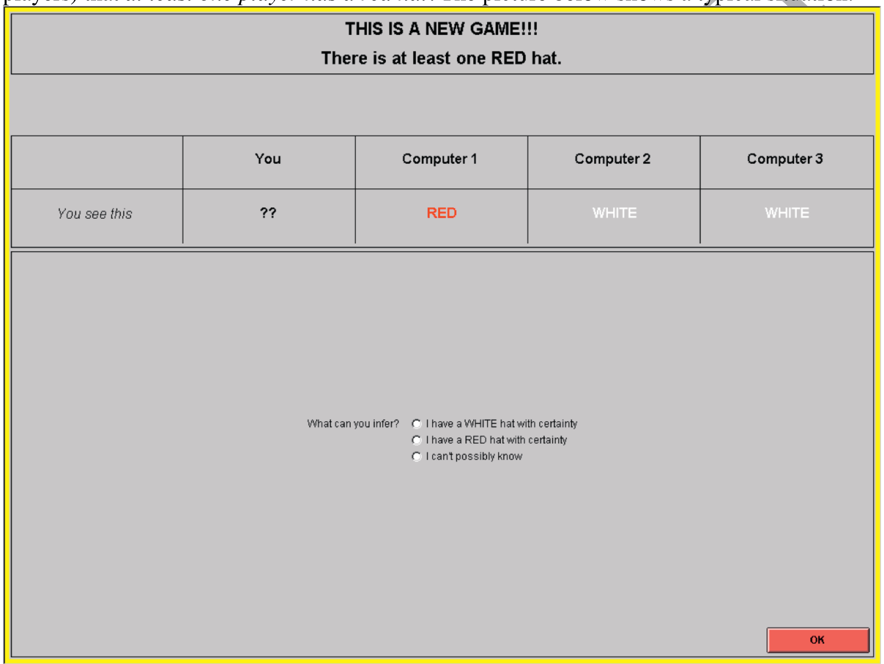
A.5 Instructions: Treatments I-IV (with repetition)

Instructions

Thank you for your participation in this experiment. If you read these instructions carefully and act upon them, you can earn real money.

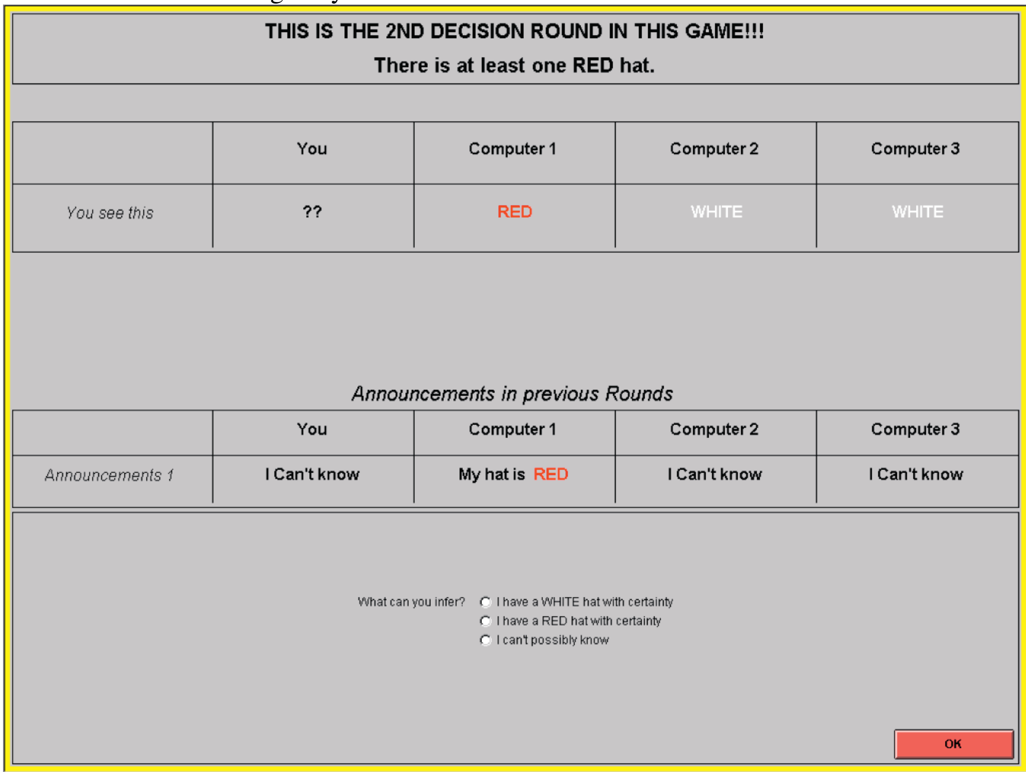
You are not allowed to communicate with other participants during the course of the experiment. If you do not follow this rule you may be excluded from the experiment.

Your task
Your task in this experiment is to determine the colour (red or white) of your hat. You will be paired with 3 computer-players. You will be able to see the colour of the hats of the computer-players, but **not** the colour of your own hat. The computer-players are in a similar situation. They observe your hat colour and the hat colours of their fellow computer-players, but not the colour of their own hat. However, everybody knows (you and the computer-players) that *at least one player has a red hat*. The picture below shows a typical situation:



You observe in this case that one of the computer-players has a red hat while the other two have white hats. The question marks "??" indicate that you do not know your hat colour.
You are asked to decide what you can infer from the information you are given. Possible answers are: "I have a WHITE hat with certainty", "I have a RED hat with certainty", and "I can't possibly know". One of these answers is correct, the two others are wrong. Note that answering "I can't possibly know" is wrong whenever it is possible to correctly infer the hat colour from the information given. Similarly ticking "I have a WHITE hat with certainty" or "I have a RED hat with certainty" is only correct if it is actually possible to logically infer that your hat colour is white or red.

The game may end after your initial decision. If the game continues, you will be given the additional information of what the computer-players have inferred from their observation. Recall that the computer-players face the same problem as you do. They can see the hats of all the others but not their own. Therefore, in the above situation, Computer 1 knows that the hats of computers' 2 and 3 are white, and also knows your hat colour. However, it does not know its own hat colour. Consequently, the computers also have a logically correct answer to the question: what can you (Computer) infer about your hat colour? The computers ALWAYS choose the logically CORRECT answer.



Above you can see a possible screen for your second decision. You again have to decide what you can infer about your hat's colour. However, now you have the additional information about what the computers (correctly) announced in the decision round before. After you have made another decision, the game may end or continue. If the game continues, you will again be given the additional information of what the computers inferred from the previous round. This process will go on until you either correctly inferred your hat colour or until you made a mistake.

<u>*Different games*</u>
After a game has ended, another new game will start. You will be given a set of 14 games.

<u>*Payment*</u>
You will start with 35 Australian Dollars. For each mistake you make, we will deduct 2.50 Dollars from your account. After the 35 games you will be paid the amount remaining in cash.

<u>*Introductory questions*</u>
Before you start the actual game we will ask you some questions about the game. These questions will be designed to test if you understand the instructions. Please make sure to read the instruction very carefully, as failing to answer the pre-game questions correctly may lead to exclusion from the experiment.

<u>*Questions*</u>
Do you have any questions? If yes please raise your hand and we will come and answer them in private.