Cognitive Forward Induction and Coordination without Common Knowledge: An Experimental Study*

Andreas Blume
Department of Economics
University of Pittsburgh
Pittsburgh, PA 15260
ablume@pitt.edu

Uri Gneezy
Rady School of Management
University of California–San Diego
La Jolla, CA 92093-0553
ugneezy@ucsd.edu

Abstract

This paper investigates optimal play in coordination games in which cognition plays an important role. In our game logically omniscient players would be able to identify a distinct coordination opportunity from other obvious facts. Real players may be unable to make the required inference. Our main experimental results are that in a coordination task with a cognitive component (1) players play differently when playing against themselves rather than against another player, and (2) given the opportunity, players signal cognition by choosing the coordination task over an outside option, a phenomenon which we refer to as *cognitive forward induction.*

Journal of Economic Literature Classification Numbers: C72, C92, D82.

* We acknowledge support through a grant from the National Science Foundation (SBR-9808947). We thank Oliver Board, Moshe Hoffman, Jack Ochs, the associate editor, an anonymous referee, seminar audiences at the ESA meetings, the Summer meetings of the Econometric Society, Arizona State University, Harvard University, Ohio State University, the University of Iowa, the University of Pittsburgh, the Technion and the WZB (Social Science Research Center, Berlin). We are especially grateful to Michael Bacharach for his detailed comments.

"Consider now a player who *is* aware of the [...] patterns."
Michael Bacharach, [*Variable Universe Games*, p. 269]

1 Introduction

In strategic interactions, cognition ability (how much a player understands the situation) is important. It is also important to understand what the player believes about the differences in cognition between herself and the other players. That is, even if she has the cognitive ability to understand the interaction, her behavior may depend on her beliefs regarding the ability of the other players. To enable us to compare the effects of cognitive ability with those of beliefs about the cognitive ability of others we employ a novel experimental design: We let subjects both play against themselves and against others. In addition, we explore to what extent experimental subjects make use of the opportunity to signal their own cognitive ability to others.

In this paper, we use *cognition* to describe the ability to arrive at new insights from readily available information. In our experiment, insights take the form of perceiving distinctions among seemingly identical objects. Cognition in our setting is closely related to computational ability, or the capacity for logical inference. In our experiment, there is a mathematically precise sense in which two obvious facts imply a third that is possibly not so obvious. Logically omniscient players automatically would make the required inference, whereas real players might fail to realize the implication.

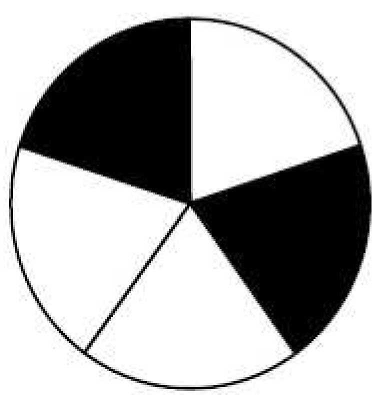


FIGURE 1

To see what we have in mind, and as a preview of the description of our experiment, consider the disk in FIGURE 1. Imagine that two players are asked to simultaneously and independently pick one of the five sectors of the disk. Both receive a positive payoff if they pick identical sectors, and nothing otherwise.

The way the disk is presented to each player, i.e. front versus back and the degree of rotation, is subject to randomization: Each of the five sectors is equally likely to be at the bottom, and "front" is as likely as "back." At the same time the integrity of the disk, i.e. the relative position of sectors is preserved. The purpose of randomizing the presentation in this manner is to leave each player with three effective choices: pick a black sector, pick one of the two white sectors that are adjacent to each other, or pick the white sector that is enclosed by the two black sectors (henceforth the *distinct sector*).

Notice that in this game, players are subject to symmetry constraints. For example, the choice "pick the black sector adjacent to the distinct sector in the clockwise direction," while available in the language of an individual player, is not part of their common language. The randomization over the presentation symmetrizes the choices of the two black sectors and of the two non-distinct white sectors.

Because of these symmetry constraints, players face *absence of a common language* in which they could label all five sectors. Crawford and Haller [1990] and Blume [2000] refer to strategies that respect such symmetry constraints as *attainable strategies.* The five-sector game has multiple Pareto–ranked equilibria in attainable strategies, e.g. "both players randomize uniformly over all sectors", or "both players randomize uniformly over all white sectors", or "both players randomize uniformly over all black sectors" (which we refer to as the *black-sector equilibrium*), or "both players pick the distinct sectors" (which we refer to as the *distinct-sector equilibrium*), etc. An optimal attainable strategy is optimal among attainable strategies. As long as the structure of the five-sector game is common knowledge, the unique optimal attainable strategy is for both players to pick the distinct sector.

In our game, however, there may be not only absence of a common language for describing choices but also absence of common knowledge of an available language due to cognitive differences. The fact that there is a unique distinct sector is a logical consequence of two other facts, first that sectors come in two distinct colors, and second that the five sectors are arranged in a circular order.[1] Even if we take the circular order and the color distinction as self-evident, players with different cognitive abilities may still come to different conclusions about whether or not there is a distinct sector. When there are such differences in cognitive abilities, there

[1]Formally the rotational and flip symmetries of the disc can be expressed through a group (in the mathematical sense). The symmetries that preserve the color distinction can likewise be expressed through a group. Combining these two structures corresponds to taking the intersection of the corresponding groups, which then reveals that there is a unique distinct sector. For details, see Blume [2000] and Blume and Gneezy [2000b].

will be players who can only describe their available choices with a *coarse language*, which only distinguishes between black and white sectors, and others who have access to a *fine language*, which in addition differentiates the distinct sector.

To analyze the game when there are cognitive differences and therefore there is absence of common knowledge of the structure of the game, we adopt the perspective of *variable frame theory (VFT)*, Bacharach [1993], expressed in terms of the symmetry constraints on strategies used by Crawford and Haller [1990] and Blume [2000]. We employ a recursive version of VFT in which we solve the game by proceeding from lower to higher cognitive levels, at each stage respecting the symmetry constraints for that level and applying the Pareto principle that was invoked by Bacharach and underlies the optimal attainable strategies of Crawford and Haller. We think of players as belonging to two cognitive levels: Low-cognition players are only aware of the black-white distinction of sectors; their strategies must respect the symmetry of the black sectors and treat the three white sectors as symmetric. High-cognition players combine the circular structure with the black-white distinction to infer that the only symmetry constraints they are facing are symmetry of the two black sectors and symmetry of those two white sectors that are adjacent to each other; they realize that they have a unique distinct choice. As a result, high–cognition players have access to a fine language in which to describe their actions, whereas low–cognition players are limited to a coarse language.

In accordance with VFT, we posit that low-cognition players do not contemplate the possibility of higher cognitive levels. Our recursive rendering of VFT then predicts that low–cognition players adopt the strategy that would be optimal if all players faced the symmetry constraints as perceived by low-cognition players; this strategy is to pick one of the black sectors (that is, low–cognition players use the optimal attainable strategy, given their perceived symmetry constraints). High-cognition players, in contrast, allow for the possibility that other players are either high– or low–cognition players. They adopt a strategy that is optimal given the symmetry constraints they are facing and given their belief about the proportion of high-cognition player. If the latter proportion is high, they pick the unique distinct sector; otherwise, they pick a black sector.[2]

If there is heterogeneity of beliefs among high-cognition players, then VFT predicts that some high cognition players will not make use of their insight that there is a distinct sector.

[2] More generally, we envision the following recursive solution: At the lowest level, L_0 players play optimally given the symmetry constraints they perceive. Higher levels L_i form beliefs about the proportions of $L_0, \dots, L_{i-1}$ players in the population, and play optimally given the symmetry constraints they perceive themselves *and* given their beliefs about the proportions and play at lower levels. This recursive solution, while in the spirit of VFT, differs from Bacharach's. His solution, translated into our terminology, is a strategy profile that respects the symmetry constraints at every level and for which there is not another strategy profile respecting the same constraints that makes players at one of the levels better off.

For high-cognition players who believe that they face a high proportion of low-cognition players it is optimal to pick a black sector. To see whether there is evidence for such behavior, we employ a novel experimental design: We compare behavior when subjects play the game against themselves with their behavior when they play against others. VFT suggests that there will be a higher incidence of choices of the distinct sector in the first case, the *self condition*, than in the latter case, the *partner condition.*

The strategic consequences of differences in object perception and their explanation in terms of VFT have been experimentally studied before by Bacharach and Bernasconi [1997] (BB). In their setting objects vary in size, shape and color. As a consequence, in BB the problem of recognizing a distinguishing characteristic (like relative smallness) is primarily a matter of whether it comes to mind or not, and not necessarily amenable to logical inference. Loosely speaking, their focus is on *psychological focal points.*

Our experiment is also concerned with object perception, except that the attributes that make one object (sector) distinct can in principle be logically inferred from two other obvious facts (black–white coloring and the five-sector circular structure). In that sense we shift the focus of the analysis in the direction of *mathematical focal points.* Thus our experiment permits us to inquire into the strategic consequences of absence of *logical omniscience,* where agents are called logically omniscient if they know all the logical consequences of their knowledge.[3] Another difference is that using a multi–sector disk that is symmetric except for color distinctions eliminates to a large degree the problem of what BB call "nuisance families," i.e. that subjects construct asymmetries (top/bottom, centrality, proximity) that were not intended by the experimenter. A third and crucial difference in our design is that we let subjects play against themselves, which permits us to measure their ability to identify the distinct sector, or more generally to measure what BB refer to as "availability" of an attribute.

Cognitive differences have also been studied in the literature on level-k reasoning in games. In its simplest form level-k theory postulates that there is a hierarchy of player types or levels, L_i, such that L_{i+1} players play a best response to L_i players, starting with an anchoring type L_0 whose behavior is exogenously specified (e.g. uniform randomization over all available actions). Sometimes, as in VFT, some fraction of equilibrium players are admitted. This approach was initiated by Nagel [1995] in the the beauty contest (a zero–sum game in which beliefs about cognition are important) and Stahl and Wilson [1994, 1995]. VFT is a close kin to level-k theory.

[3] Lipman [1999] relates absence of logical omniscience to framing effects. In our case framing effects are relevant for three reasons: First, the existence of a distinct sector must be inferred from other facts; second, different players may have access to different frames (or symmetry constraints); and third, we manipulate frames by considering both separated black sectors as in FIGURE 1 and adjacent black sectors.

The principal difference is in focus. VFT and the present study concentrate on differences in the perceived description of the game as opposed to differences in the depth of reasoning about best responses.

Bacharach and Stahl [2000] develop a boundedly rational version of VFT that has room for both types of belief hierarchies, those due to differences in the perceived description of the game and and those arising from differences in depth of reasoning. Bacharach and Stahl allow symmetry constraints to vary for the anchoring type, L_0. In our setting, this means that there could be an L_0 player who distinguishes only black and white sectors, but also an L_0 player who in addition can identify the distinct sector. The defining characteristic of an L_0 player according to Bacharach and Stahl is that he randomizes uniformly over classes of actions, where each class consists of actions that are symmetric to each other and is maximal in that respect. In our game, a coarse–language L_0 player would assign probability one–half each to picking a black sector and picking a white sector. A fine–language L_0 player would assign probability one–third each to picking a black sector, picking the distinct sector, and picking a non–distinct white sector. Notice that already among L_0 players there is a tendency to favor black sectors and the distinct sector over the other two white sector. In the Bacharach–Stahl setup the details of L_1 behavior are complex, depending on his own symmetry constraints, which can be more or less severe than those of an L_0 player, his beliefs about the symmetry constraints of L_0 players etc. For our purposes it suffices to allow only one type of L_1 player, one who has a fine language, forms beliefs about the frequency of fine– versus coarse–language L_0 player, and best responds to those beliefs. Importantly, just as under VFT such an L_1 player will pick the distinct sector if he strongly believes that the L_0 player has a fine language and will pick a black sector otherwise.[4]

VFT, being designed for one-shot games, is silent about the possibility that in a dynamic setting players may come to have an insight during the course of the game that initially eludes them. This effect potentially contaminates a within-subject comparison of the self condition and the partner condition. Therefore, while we include a within-subject comparison, we believe that the real test of VFT is in our between-subject comparison.

VFT is equally silent about the possibility that in a dynamic setting players may try to make inferences about the cognition of others or signal their own cognition. Nevertheless, if we follow the basic premise of VFT that there are different cognitive levels and that high–cognition players reason about the relative frequency of high– and low–cognition players in the population,

[4]Belief hierarchies also naturally appear in models with behavioral types. Crawford [2003], for example, has studied deception in communication games in which fully rational players interact with boundedly rational ones. Experimental evidence for deception and credulity has been reported in Forsythe, Lundholm and Rietz [1999] and Blume, DeJong, Kim and Sprinkle [2001].

VFT suggests the possibility of what we call *cognitive forward induction.* In standard game theory, a *forward induction* argument can sometimes be used to question the stability of a Nash equilibrium in the game. This occurs, for example, when a player has a deviation from his equilibrium strategy that leads to a continuation game in which there is a unique equilibrium that yields him a higher payoff than his status-quo equilibrium payoff. One can then imagine the player accompanying his deviation by the speech: *We both know the payoff I would have received if I had stuck to my equilibrium strategy. Since you know that I am rational you should infer that there was a purpose to my deviation and that it must be that I am aiming for the unique Nash equilibrium in the continuation game that would raise my payoff relative to the status quo.*

We set up a similar scenario in our experiment. One player (or possibly both) is given the choice between an outside option and playing the five-sector game described above. Payoffs in the experiment satisfy two requirement: First, the expected outside-option payoff is no less than the expected payoff from the black-sector equilibrium in the five-sector game. Second the payoff from the distinct-sector equilibrium exceeds the expected payoff from the outside option. A high-cognition player may then consider choosing the five-sector game and hope to have his choice interpreted as in the following speech: *We both know the payoff I would have received if I had taken the outside option. Since you know that I am rational you should infer that there was a purpose to my choice and that it must be to convince you that I understand that we have a unique distinct choice and that if both of us make that choice I would raise my payoff above the outside–option payoff.* According to this reasoning, we would then expect that players choosing to play the five-sector game are more likely to pick the distinct sector than players in our partner condition. Furthermore, we should expect that players who are told that their parter had a choice of which game to play show a higher incidence of distinct sector choices than players under the baseline partner condition.

Our main experimental results provide some support for these predictions. We find evidence that (1) in the five-sector game players may fail to coordinate with themselves, (2) in the five-sector game players play differently when playing against themselves rather than against another player, (3) given the opportunity, players will signal cognition by choosing the five sector game over an outside option, and (4) players who are conscripted by other players into playing the five-sector game are more likely to pick the distinct sector than players in the five-sector baseline game (where there is no choice).

The inability of some players to solve the game when playing against themselves suggests there are indeed cognitive difference with respect to the experimental task. The difference in behavior when playing against another player suggests that the coordination failures we observe are as much a consequence of a failure of common knowledge of cognition as they are a failure of cognition itself. Cognitively more astute players appear to be aware of differences in cognition. Signaling of cognition provides further evidence that high cognition players believe that other high cognition players form beliefs about cognition.

The paper is organized as follows. The next section describes the experimental design. In Section 3 we discuss predictions for our games that are inspired by VFT. In section 4 we report and discuss the experimental results. Section 5 discusses the literature and Section 6 concludes.

2 Experimental Design

In each of our experimental sessions participants made choices on a pie chart that was divided into five sectors of equal size, two of which were black and three of which were white. We considered two configurations of the pie charts. In one, the two black sectors where *separated* from each other, as in FIGURE 1, and in the other they were *adjacent* to each other. In either case there is a unique distinct sector; when the two black sectors are separated from each other it is the white sector that is enclosed by the two black sectors, when the two black sectors are adjacent to each other it is the white sector that is opposite from the two black sectors.

We had eight experimental sessions. Each session was conducted with a different cohort of participants. Each of the first five cohorts had 30 members, the sixth cohort had 100 members, the seventh had 154 members and and the eighth had 160 members. In our first five sessions we compared how subjects played against themselves with how they played against others. In these sessions we also varied the configuration of black sectors. These five sessions correspond to four basic treatments, and a variation of one of the basic treatments. The sixth, seventh and eighth session correspond to distinct forward induction treatments, in which we let (some) subjects choose which game to play. We first describe the procedure for the base game that was implemented in the *Partner-Separated Session*, in which the two black sectors were separated by a white sector as shown in FIGURE 1, each subject was randomly matched with another subject, and there was no choice of games. Then we will describe how the other treatments differed.

We invited students into a room and gave each a registration number. The students then received written instructions, which were read aloud to make them common knowledge. They were told that only the investigator would know the identity of the person with whom they were matched.

A pie chart, consisting of a separate front and back, with five equal-sized sectors, three white, two black, and the black sectors separated (as in FIGURE 1) was then shown to the subjects. The back and front of the pie chart looked identical. They were glued together to leave space for marking choices. Each participant was asked to choose a sector. Using a sticker, we then marked this choice on the inside, invisible from either side. With probability one-half, the investigator turned over the chart, exchanging front and back. The matched participant was then asked (without knowing the choice of his counterpart) to choose a sector. Participants were told that if the sectors chosen matched, each would receive a positive payoff, and otherwise nothing. The whole procedure was then repeated in a second stage, without revealing the outcome from the first stage.

The two-by-two design for the first four treatments is described in TABLE I.

Table I

Two-by-Two Experimental Design

	Separated black sectors	Adjacent black sectors
Partner	Partner–Separated	Partner–Adjacent
Self	Self–Separated	Self–Adjacent

The difference between the Partner–Separated treatment, which was described above and the Partner–Adjacent treatment concerns the configuration of the black sectors. In the Partner–Adjacent treatment, the two black sectors are adjacent to each other. The Self–Separated and the Self–Adjacent treatment were the same as the other two treatments, except that participants were matched with themselves: Participants made two choices. In stage one each participant was asked to choose one of the five sectors. The experimenter recorded that choice with a sticker between the two layers of the disk. The disk was then randomly rotated and randomly flipped before being presented again to the same participant. In stage two each participant was again asked to choose one of the five sectors on the disk.

We ran two variations of the Partner–Separated treatment, one in which everyone played against a partner twice, without feedback, and one in which everyone first played against a partner twice and then against him- or herself, again without feedback between stages. Also the Self–Separated treatment was succeeded by a single stage of play against a partner, without prior feedback. The Partner–Adjacent treatment involved three repetitions of play against a partner, without feedback between stages. The Self–Adjacent treatment consisted of two stages of playing against oneself followed by one stage of playing against a partner, once more without feedback between stages.

Our last three sessions corresponded to the forward induction (FI) treatments. In Session 6, all participants were given the choice of either playing a two-sector game, with two white sectors, or the five-sector base game with two separated black sectors and three white sectors. Anyone choosing the two-sector game played the two-sector game. Whoever chose the five-sector game played the five-sector game, provided there was another participant who had made the same choice. That way five-sector play was only by choice.

In Session 7 participants were divided into two equal-sized groups of *choosers* and *candidates.* Choosers had the opportunity to decide between playing the five-sector game (with two separated black sectors) and an outside option. Each chooser who opted in was matched with one of the candidates to play the five-sector game. Candidates who were matched with choosers became *conscripts*, who had to play the five-sector game. The roles of choosers and candidates was common knowledge. For candidates, we employed the strategy method, that is they had to specify their choices in the five-sector game both for the event of being conscripted by their partner and the event that their parter chose the outside option.

Session 8 was identical to Session 7 except that conscripts did not learn that their partners had a choice between playing the five-sector game and an outside option. Choosers were informed that candidates would not learn of their ability to decide between playing the game and taking the outside option. Candidates who were not conscripted were matched with other such candidates to play the five-sector game.

The first six sessions were conducted at the Technion. Participants earned 20 Shekels (at that time 1 Shekel was equivalent to $0.26) for each coordinated choice.[5] Sessions 7 and 8 were run at the University of California–San Diego. In these sessions participants earned $10.00 for each coordinated choice and the outside option was worth $6.00.

3 Predictions

A natural benchmark with which to compare our experimental data is the prescription that a social planner who is aware of the players' symmetry constraints would make to them. In the five-sector game a social planner would solve a constrained optimization problem, maximizing the joint payoff subject to the symmetry constraints on strategies that are implied by the experimental design. The solution to this problem is to recommend to pick the distinct sector. When the two black sectors are separated, the distinct sector is the one enclosed by the two black sector. When the two black sectors are adjacent to each other, the distinct sector is the one opposite of the two black sectors. Picking the distinct sector is the *optimal attainable strategy,*

[5]Translations of the instructions are in the appendix.

OAS, of Crawford and Haller [1990] and Blume [2000].

While useful as a benchmark, there are reasons to expect that OAS cannot account for all of our data. One is simply noise in the data and could be easily accommodated. Another is that prior work suggests that there are systematic deviations from the OAS prediction in similar games. For example, Blume and Gneezy [2000a] (BG2000) find that with a nine-sector disk and two marked sectors, even though there is also a unique distinct sector, it was *never* chosen by experimental subjects.

One possible explanation for the results of BG2000 is that experimental subjects are simply unable to identify the distinct sector. It is also possible, however, that while there are at least some who identify the distinct sector, they are not sufficiently confident that others identify it as well. The latter explanation can be formalized by combining OAS reasoning, i.e. optimization subject to symmetry constraints, with cognitive hierarchies, i.e. allowing players to perceive symmetry constraints differently.

The result is (a generalization of) Bacharach's *variable frame theory (VFT).*[6] For our game the theory postulates that there is a hierarchy of cognitive types. High-cognition players are able to identify the distinct sectors; they are only subject to the symmetry constraints (between the two black sector, and between the two non–distinct white sectors) that the environment imposes on them. Low–cognition players are subject to additional symmetry constraints; their strategies must treat all three white sectors as identical.

VFT further postulates that low-cognition players do not contemplate the presence of higher cognitive levels; they adopt a strategy that is jointly optimal for low–cognition players given the symmetry constraints they perceive. This strategy is to pick a black sector. High–cognition players form beliefs about the relative likelihood of low– and high–cognition players. They adopt a strategy that is jointly optimal for high–cognition players, given their beliefs. If they believe that there is a high probability of facing another high-cognition player, they pick the distinct sector. Otherwise, it is optimal for them to ignore their insight and to pick a black sector.

We will also postulate that there is heterogeneity of these beliefs in the population, and that the distribution of beliefs has full support on the unit interval. As a result, VFT predicts that there is a higher incidence of picking the distinct sector, when players play against themselves

[6]Bacharach's formulation of VFT requires that players' strategy spaces be redefined in terms of "acts." Acts group together actions that are symmetric according to a player's "frame." Our formulation, using symmetry constraints rather than "frames," generalizes VFT in a way that easily accommodates properties like adjacency, centrality etc, does not require us to change player's strategy spaces, and gives us immediate access to the entire arsenal of group theory to let us express different symmetries. For the game at hand, this means that we have a simple way of thinking about how the circular structure of the disk interacts with the black-white structure: The intersection of the corresponding groups formally identifies the distinct sector. Casajus [2001] offers an alternative generalization of VFT.

than when they play against others. This is summarized in the following hypothesis:

Hypothesis 1 [Beliefs Matter] *The observed fraction of choices of the distinct white sector when players play against others is less than the observed fraction of choices of the distinct white sector when they play against themselves.*

The hypothesis expresses the idea that high-cognition players, those who manage to identify the distinct sector, form beliefs about the relative proportions of high– and low–cognition players in the population. A word of caution is in order here regarding our ability to measure whether a given player has high or low cognition. Having a player play against himself may trigger an insight that switches a player from low to high cognition. There may be an *uncertainly principle* at work here in that we cannot measure a player's cognition without altering it. For that reason, while we provide within-subject comparisons, we believe that the proper way of testing the Beliefs-Matter hypothesis is via between-subject comparisons. The within-subject comparisons then provide a way of evaluating the concern of whether indeed there is an uncertainty principle (that we cannot simultaneous measure a players cognition and his belief about other players) at work here.

While VFT was formulated for one-shot games, its spirit straightforwardly extends to dynamic settings and raises intriguing questions. Differences in cognition may be affected by the history of play, beliefs about cognitive differences may change as a function of history, and players may have an incentive to signal their own cognition. We refer to players trying to signal cognition as *cognitive forward induction.* To see whether there is evidence for cognitive forward induction, we give (some) players a choice between the five-sector game and an outside option whose expected payoff is no higher than the payoff from the strategy of picking a black sector in the in the five-sector game, and in two of our treatments we make it public knowledge that such a choice opportunity exists.[7] If cognitive–forward induction considerations affect behavior,

[7] In two of our three forward–induction treatments the outside–option payoff is strictly between the payoffs from the black–sector equilibrium and the distinct–sector equilibrium of the five–sector game. Under these conditions, even in the absence of cognitive issues the conventional forward–induction arguments apply: The strategy to enter the five–sector game and then to select a black sector is strictly dominated by choosing the outside option. Similarly, the strategy to enter the five–sector game and then to pick one of the two non–distinct white sectors is strictly dominated. Hence, in the absence of cognitive constraints players who find themselves in the five–sector game should infer that their partner must intend to pick the distinct sector. Hence, all players should choose the five sector game. Once there are cognitive constraints, we get the additional effect that coarse–language players who have a choice between games should choose the outside option, which give them a strictly higher expected payoff than the highest expected payoff they can hope for in the five–sector game. Furthermore, even if as in one of our treatments, the outside option payoff does not exceed but is equal to the black–sector equilibrium payoff, for a coarse–language player picking the outside option weakly dominates entering the game: His expected payoff from entering the game is at most equal to his expected payoff from the outside option, but he risks that the other player does not also pick a black sector. For a coarse–language player the expected payoff when the other player picks a white sector is at most two thirds of the expected outside option payoff.

then (at least for some players) choice of the five–sector game is a signal of intent to pick the distinct sector and (at least some of the) players who are drafted into playing the five–sector game interpret the signal correctly, provided they know that their partners had a choice of which game to play. Accordingly, our second major hypothesis is the forward induction hypothesis:

Hypothesis 2 [Forward Induction]

1. *Players who choose the five-sector game over the outside option are more likely to pick the distinct sector than players in the five-sector game when there is no opportunity for choice.*
2. *Players who are conscripted to play the five-sector game likewise are more likely to pick the distinct sector than players in the five-sector game when there is no opportunity for choice.*
3. *The five sector game is more frequently chosen when the choice opportunity is publicly known than when only choosers are aware of the choice opportunity.*

By choosing the five sector game, high-cognition players effectively send the message "I must be a high-cognition player because only for high-cognition players is there a possible advantage from choosing the five-sector game." Players who observe their partner having chosen the five–sector game may be affected in one of two ways, depending on their cognitive type. High-cognition players may infer that their partner is a high-cognition player as well; low-cognition players, if they pay any attention to the signal at all, may be prompted to reexamine their perception of the game.

The distinctive characteristic of VFT is that players form beliefs about the cognitive types of other players and follow equilibrium strategies given their beliefs. By dropping either of these two components we obtain two natural hypotheses that compete with VFT. One competing hypothesis is that players are naive and believe that others face the same constraints as themselves. Under this *naive hypothesis* there would be no difference between playing against oneself and playing against others. High-cognition players would pick the distinct sector and low-cognition players one of the black sectors. Furthermore, in contrast to our forward–induction hypothesis the *naive hypothesis* would predict that players who have been conscripted to play the five–sector game by other players behave no differently than they would if they simply had to play the game in an environment without choice of games.

An alternative hypothesis competing with rational VFT results when we drop the equilibrium component. This gives us a version of *level–k theory*, in which players at level j in the cognitive hierarchy, L_j players, best respond to L_{j-1} players. As discussed in the introduction, Bacharach and Stahl [2000] propose a boundedly rational version of variable frame theory along these lines. The predictions of level–k theory depend on the choice of the anchoring type L_0. If we follow Bacharach and Stahl to allow the symmetry constraints to vary for the anchoring type (as discussed in the introduction), the predictions from boundedly rational VFT do not substantially differ from rational VFT. One minor difference is that boundedly rational VFT can readily account for choices of non–distinct white sectors in the data.

In the terminology of Bacharach and Bernasconi [1997], our disk game is a *trade–off game*: There is a tension between the obviousness of the fact that there are fewer black than white sectors and the uniqueness of the distinct sector. For such games Bacharach and Bernasconi formulate their trade-off theorem, which in our game amounts to: "Players who recognize the distinct sector pick that sector if and only if the probability of others recognizing it exceeds the frequency of the black sectors. Otherwise they pick a black sector." This makes sense if high–cognition players have accurate and identical estimates of the frequency of low–cognition players in the populations. It is more likely, however, that these estimates vary among high–cognition players. Therefore we suggest testing a weaker TRADE-OFF HYPOTHESIS:

Hypothesis 3 [TRADE-OFF] *If the relative frequency of distinct–sector choices and black–sector choices varies in the self–condition between two different five–sector games, it varies in the same direction in the partner–condition.*

To investigate the trade–off hypothesis, we compare two versions of our five–sector game. In one version the two black sectors are separated by the distinct white sector, as in FIGURE 1; in the other version the two black sectors are adjacent to each other, which induces a distinct sector that is opposite to the two black sectors. This comparison was motivated by the hunch (confirmed in the data, as we will see) that participants would find it more difficult to identify the distinct sector, when the black sectors are adjacent to each other.

Recently, there has been increasing interest among economic theorists in modelling decision problems and games when agents may be unaware of some facts, e.g. Dekel, Lipman, Rustichini [1998], Fagin and Halpern [1988], Geanakoplos [1989], Modica and Rustichini [1999], Halpern and Rêgo [2005], Heifetz, Meier and Schipper [2006] and Board and Chung [2007]. The aim of this literature is to provide a language for incorporating unawareness into models of decisions and games. Some desiderata for such a language are that it permits one to express beliefs about awareness and mutual awareness; more recent contributions have also attempted to express agents' beliefs about (their own) unawareness.

VFT formulates a hypothesis of how players deal with a specific type of unawareness, i.e. about degrees of symmetry, in a specific class of games, i.e. pure coordination games. From this perspective our experiment is designed to create an environment in which we can demonstrate the existence of unawareness of a fact (the existence of a distinct sector) among a fraction of subjects and investigate to what degree experimental subjects reason about the unawareness of others. The principles that guide VFT are similar to the ones used by Heifetz, Meier and Schipper [2007] in their formulation of Bayesian games with unawareness and their proposal of an equilibrium concept for such games: They introduce unawareness types, which are ranked by the richness of the vocabulary with which they can describe the world. This is analogous to our "coarse language–fine language" distinction. An important part of their equilibrium construction is what they refer to as the "tyranny of the unaware:" awareness types with narrow horizons do not conceive of awareness types with wider horizons. This corresponds to our assumption that coarse–language players do not envision players with a finer language, what Bacharach and Bernasconi refer to as "blinkering" of players. In their words "... because of blinkering, [Player 1] does not in general have a complete list of [Player 2s] possible types." Equilibria in the formulation of Heifetz, Meier and Schipper extend from lower to higher levels of awareness, which is analogous to our recursive construction of equilibrium.

Games with incomplete awareness are also studied by Feinberg [2005]. He is interested in modelling unawareness of other players' actions, which fits in nicely with our setting in which coarse–language players fail to recognize that there is a choice of picking the distinct sector. Note that in our game a player may be unaware of the availability of some of his own actions.

4 Experimental Results

The results from our Partner–Separated treatment are summarized in Figure 2. In this and the following figures and tables, "*d*" denotes the choice of the distinct sector, "*b*" the choice of a black sector, and "*w*" the choice of a white sector other than the distinct sector. In Figure 2 we report frequencies of sequences of choices for the two stages of the Partner-Separated treatment. Thus, for example *bd* denotes a sequence in which a subject chose a black sector in the first period and the distinct sector in the second period.

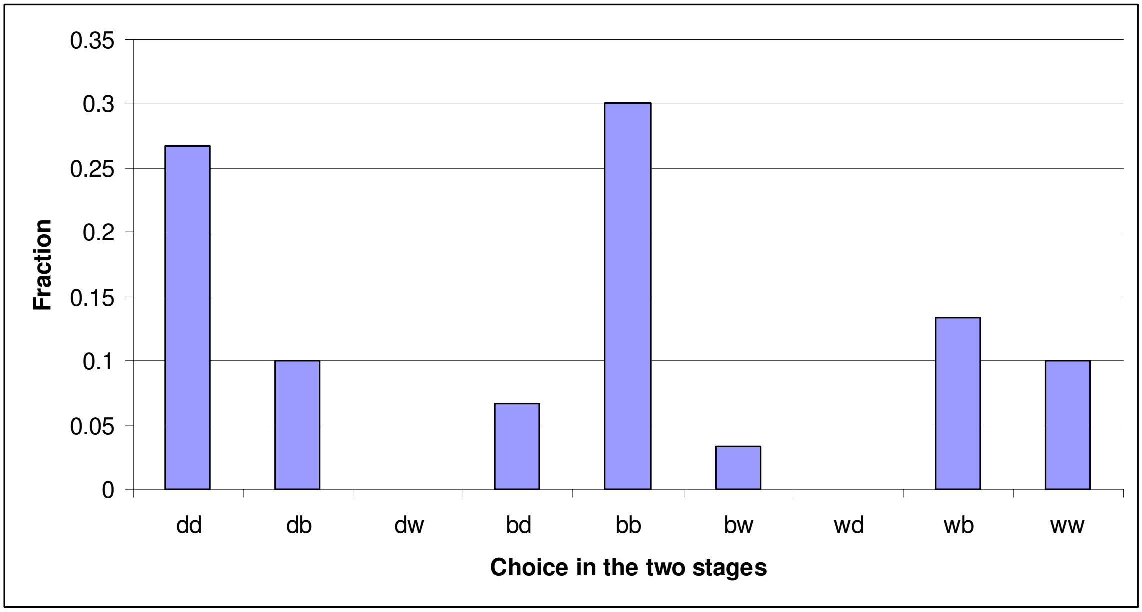


FIGURE 2 The Partner–Separated Treatment

The distribution has two modes, one where subjects consistently chose one of the black sectors, the other where subjects consistently chose the distinct sector. Like in Blume and Gneezy [2000a], subjects systematically deviate from the full-information OAS prediction that they pick the distinct sector.

There is a preponderance of d and b choices. Indeed, the frequency of "w" choices appears less than chance would predict; in the first stage, we can reject the hypothesis that the probability of "w" is 2/5 (the p-value is .044); in the second stage, we can reject the same hypothesis (the p-value is .002).

Figure 2 suggests that there was a fair degree of persistence between stages. In 16 out of 25 cases the subjects repeated their first-stage choice in the second stage. If choices were random, we would expect less than 2/5 of the first-stage choice to be repeated in the second stage. We can reject the hypothesis of no persistence, $p = .003$

It is also worth noting that the situation that subjects face in the Partner-Separated treatment is essentially the same as the situation of the 80 candidates in the Forward Induction treatment with *privately* known role assignments that we will discuss in detail later. In this treatment the candidates did not know that choosers could select the game and candidates always played the five-sector game. Of the 80 candidates 18 picked the distinct sector and 56 picked one of the black sectors.

The results from our Self-Separated treatment are summarized in FIGURE 3.

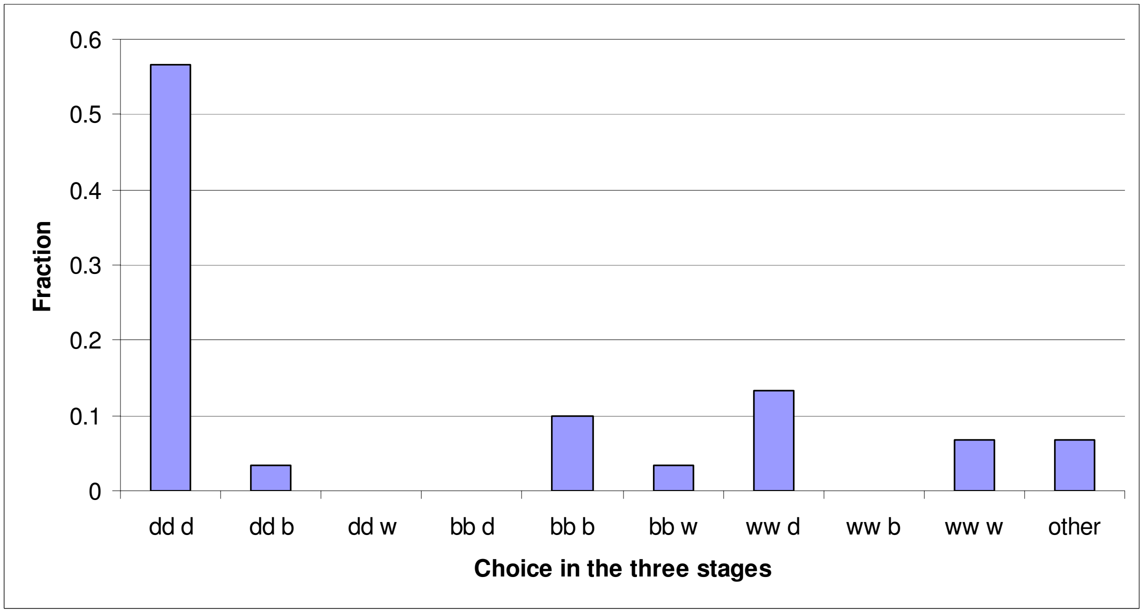


Figure 3 The Self-Separated Treatment

The first thing to note is that when players play against themselves there is a marked increase in the choice of the distinct sector. The fraction of d choices in the first stage of the Self-Separated treatment is a measure ν of the frequency of high-cognition player. Consistent with our Beliefs Matter hypothesis, the fraction of d choices in the first stage of the game with a partner, ϕ, satisfies $\phi < \nu$. The p-value for a test of equality of proportions is $p = .06$.[8] If we combine the data from the first stage of the Partner-Separated Treatment with those for the 80 candidates in the Forward Induction treatment with *privately* known role assignments, the p–value becomes $p = 0.00045$. Therefore, we find strong support for our key hypothesis that players' choices are not only determined by what they know but also by what they believe.

The second remarkable fact is that after having played against themselves, subjects are twice as likely to choose the distinct sector, when paired with a partner. Formally, we can reject the hypothesis that the proportion of distinct sector play is the same with and without prior play against oneself ($p < .001$). This is not explained by conventional learning theories since between rounds there is no feedback on either own payoffs, or the choices of other players. This observation is consistent with what we referred to as the *uncertainty principle* that measuring the fraction of high-cognition players affects the measured proportion of distinct sector play.

If prior play against oneself has such a dramatic effect on play against a partner, one may wonder whether the reverse is true as well. It could be that in either case the task becomes

[8]Here and in similar test below, we use the one-tailed p-value from Fisher's exact test.

easier by prior exposure to a similar task. However, depending on the task performed first, it is also possible that this task creates a frame that adversely affects performance in the second task. To answer this question, we conducted a variation of our first *(Partner-Separated)* treatment, in which we first let the participants play against a partner and then against themselves. The data from this treatment are reported in FIGURE 4.

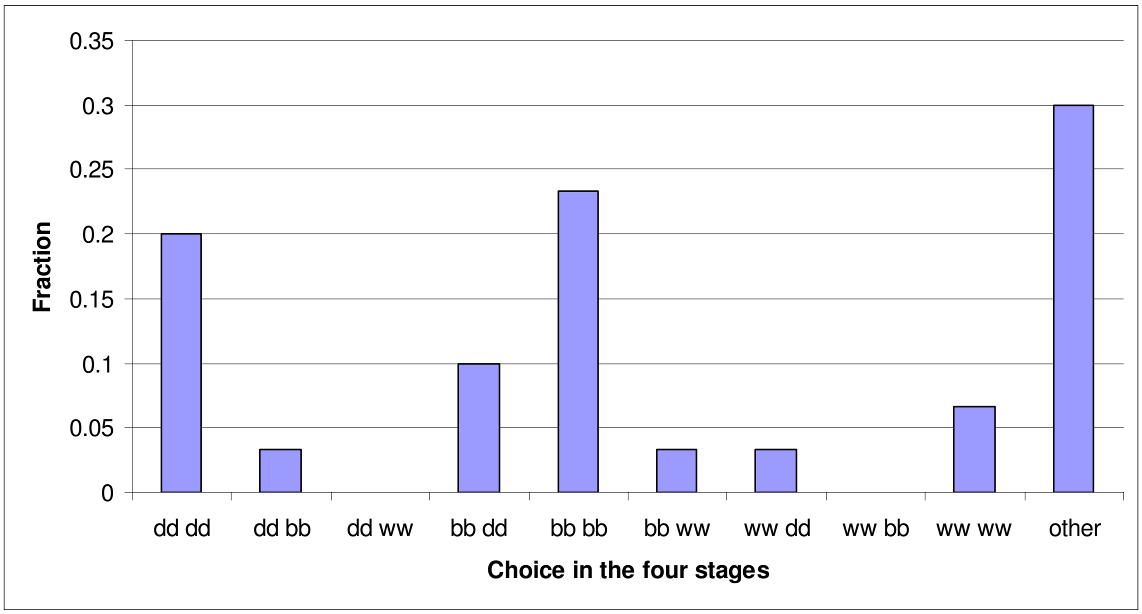


FIGURE 4 Partner-Separated followed by Self-Separated Treatment

First, note that the results reported for the first two stages in FIGURE 4 (against a partner) are very similar to those reported for the two stages in FIGURE 2. This confirms and strengthens our results on play against a partner and provides further evidence for our BELIEFS MATTER hypothesis. In contrast, the frequency of distinct-sector play in stages 3 and 4 of FIGURE 4 (against oneself) is slightly lower than in the first two stages of the self-prominent treatment (also against oneself) in FIGURE 3. Most importantly, the frequency of d choices when playing against oneself after playing against a partner (30 out of 60 in stages 3 and 4 in Figure 4) is considerably lower than the frequency of d choices when playing against a partner after first playing against oneself (24 out of 30). Altogether, prior play against oneself appears to have a salutary effect on subsequent play, whereas prior play against a partner does not have such an effect.

While VFT has been formulated for static settings, it is not difficult to come up with amendments that would explain the salutary effect of playing against oneself for subsequent play against others. Players on their own are not as likely to "put themselves into the other player's shoes" as when they are forced to do so by having to assume both roles. Thus, if they first play against themselves they may learn to better to understand the strategic situation and can use the knowledge gained this way in later play against others. Analyzing the game from both perspectives may lead some individuals to discover that there is a unique distinct sector. In contrast, the observation that prior play against a partner seemingly adversely affects cognition is puzzling from a VFT perspective. How can it be that any kind of exposure to the game appears to adversely affect cognition? Perhaps, when first playing against a partner, a player may get locked into a disadvantageous strategy.

The results from the Partner-Adjacent treatment are summarized in FIGURE 5.

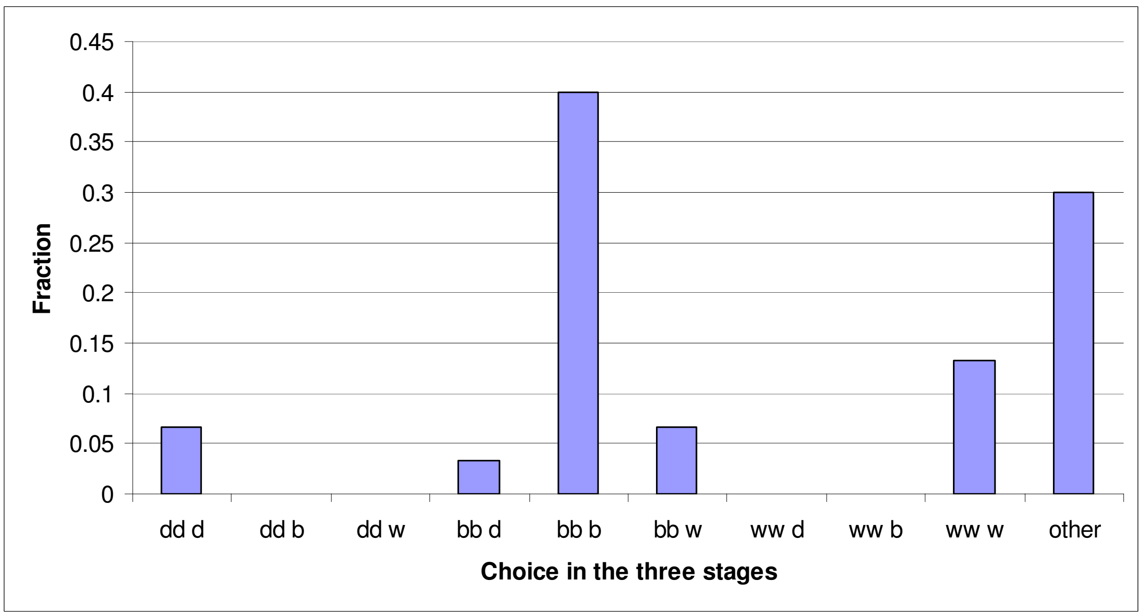


FIGURE 5 The Partner-Adjacent Treatment

In the Partner-Adjacent treatment we observe a marked drop in distinct sector play relative to the Partner–Separated treatment. Formally, we can reject the hypothesis of equality of the proportions in the Partner-Separated and Partner-Adjacent treatments ($p < .036$). In the Partner-Adjacent treatment the modal choice in the first stage is to pick a black sector. There were 18 choices of a black sector in the first-stage, while we would expect only 12 on average, if choices were made randomly. Indeed, we can reject the randomness hypothesis, $p = .021$.

The results from the Self-Adjacent treatment are summarized in FIGURE 6.

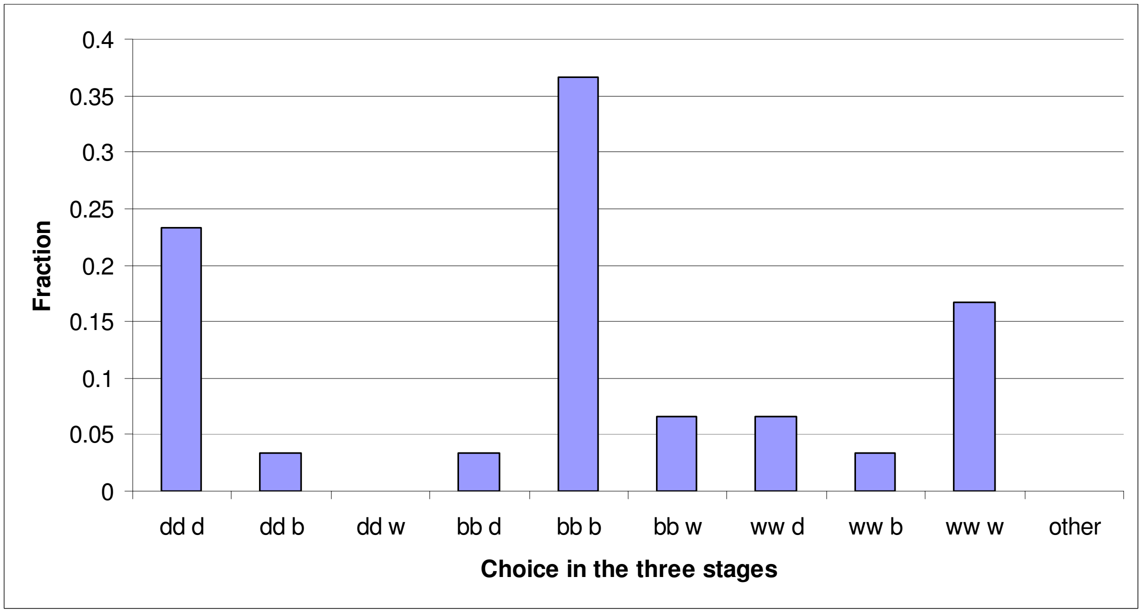


FIGURE 6 The Self-Adjacent Treatment

As predicted by our BELIEFS MATTER hypothesis, the observed fraction of choices of the distinct sector, ϕ, in the Partner-Adjacent treatment is less than the fraction of high–cognition players, ν as measured by the self-adjacent treatment, although the difference is not statistically significant, $p = .167$. As in the separated-sector treatments, prior play against oneself does seem to encourage play of the distinct sector, when later paired with a partner, consistent with our uncertainty principle, although here the evidence is weaker, $p = .063$.

The adjacent–sector and separated–sector treatments represent two ways of framing essentially identical problems. This permits us to ask how our TRADE-OFF HYPOTHESIS fares in the comparison between the two frames. We find that a smaller fraction of participants picks the distinct sector in the first stage of the Self–Adjacent treatment (8 out of 30) than in the first stage of the Self–Separated treatment (17 out of 30), $p = .0176$. Our TRADE-OFF hypothesis therefore predicts that a smaller fraction will pick the distinct sector in the Partner–Adjacent treatment than in the Partner–Separated treatment. Indeed, confirming the hypothesis, the proportions are 4 out of 30 versus 11 out of 30 for the first stages of the two treatments respectively, $p = .0358$. Notice that while the naïve hypothesis would make the same directional prediction as our TRADE-OFF hypothesis, the naïve hypothesis also predicts that there is no difference between the self and the partner conditions.

Recall that we evaluate our FORWARD INDUCTION hypothesis with three different treatments. In our first forward-induction treatment all participants are given the choice between playing the five-sector game with separated black sectors or a two-sector game. As shown in Figure 7, most participants chose the outside option and conditional on choosing the five–sector game a large majority picked the distinct sector.

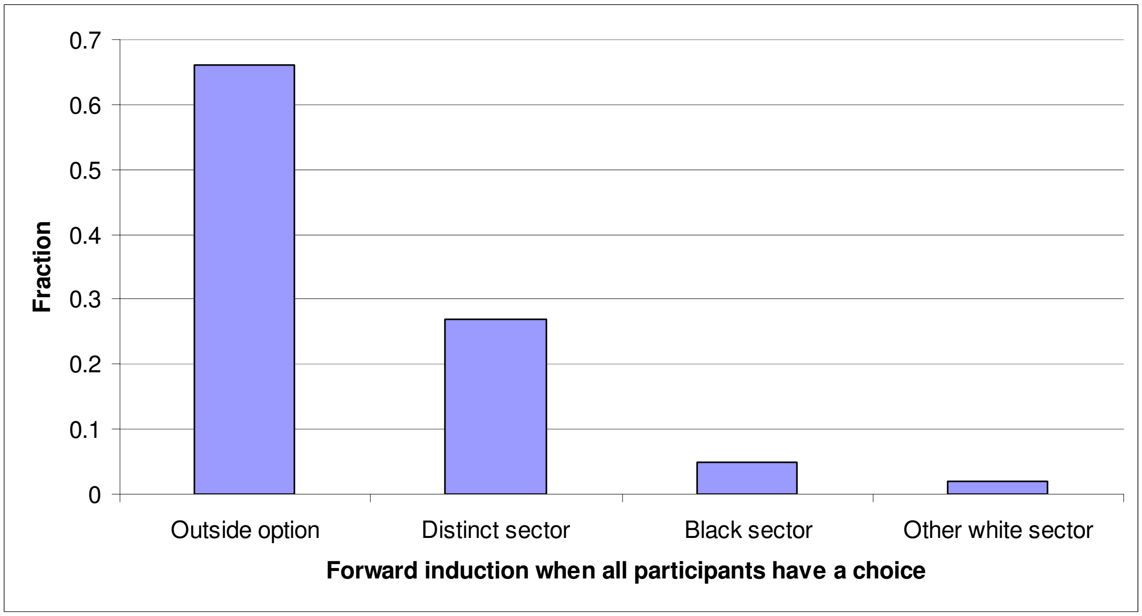


FIGURE 7 Forward Induction - All Participants Have a Choice

In this treatment, 34 out of 100 participants chose the five-sector game. 27 out of the 34 proceeded to choose the distinct sector. 5 chose one of the black sectors, and two chose one of the other white sectors. This implies that there is a marked increase in the percentage of distinct sector play, when we allow players to select the game rather than have everyone play the game, $p < .001$. This is consistent with (the first part of) our FORWARD INDUCTION hypothesis.

Giving all participants a choice, and having the expected payoff from the two-sector outside option be the same as from both players picking the black sector does not result in the cleanest possible test of our Forward–Induction hypothesis. First, the payoff equivalence may lead some low-cognition players to choose the five-sector game. Second, it could be that even under the naïve hypothesis sufficiently optimistic high-cognition players enter the five-sector game and then choose the distinct sector, without making inferences about why their partner may have chosen the five-sector game.

For that reason we considered a second forward-induction treatment in which different players are assigned different roles. Some players, *choosers*, have a choice between an outside option and playing the five-sector game. Other players, *candidates*, only play the five-sector game if they are conscripted by their partner to play. The payoff from the outside option exceeds the expected payoff from both players picking a black sector in the five-sector game. This structure, including the role assignment, is public knowledge. Therefore, we call this treatment Forward-Induction with *publicly* known role assignments. The results from this treatment are summarized in FIGURE 8.

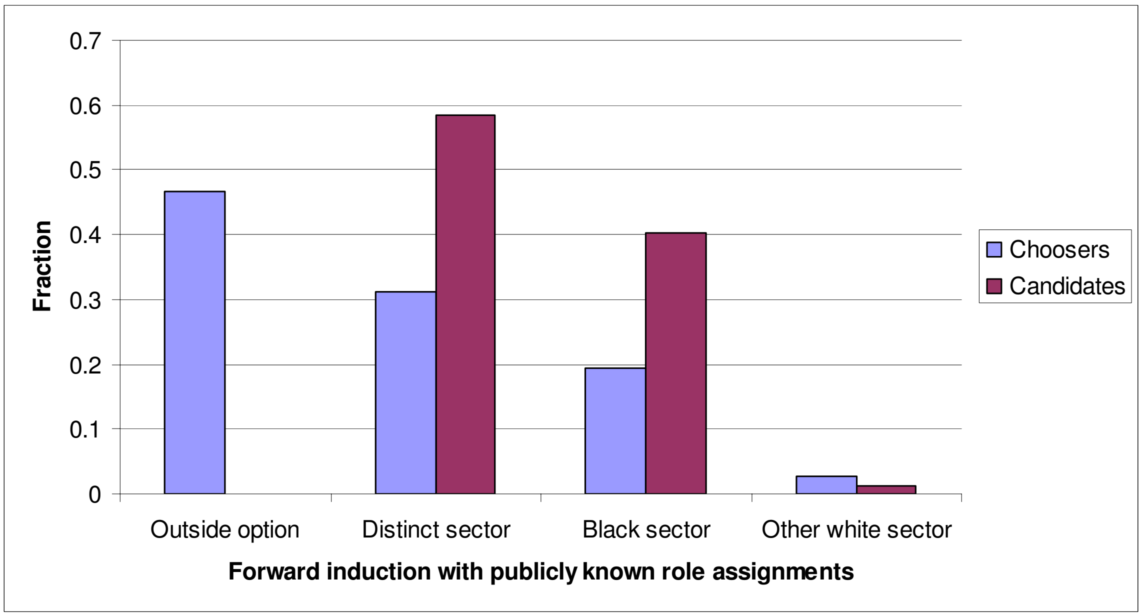


FIGURE 8 Forward Induction - Publicly Known Role Assignments

24 of the 41 choosers who elected to play the five-sector game picked the distinct sector (58.5%). In contrast, only 11 out of 30 (36.6%) participants picked the distinct sector in the first stage of the Partner-Separated treatment ($p = 0.057$); if we combine these data with those from the 80 candidates in the Forward Induction treatment with *privately* known role assignments, where 18 out of 80 candidates chose the distinct sector, the p–value becomes $p = 0.0003$. These results lend strong support to the first component of our FORWARD INDUCTION hypothesis; choosers who elect to play the five-sector game are more likely to pick the distinct sector than subjects in the five-sector game without choice.

41 out of 77 candidates picked the distinct sector (53.2%) in contrast to 11 out of 30 in the first stage of the Partner-Separated treatment ($p = 0.09$); if we combine these data with those from the 80 candidates in the Forward Induction treatment with *privately* known role assignments, where 18 out of 80 candidates chose the distinct sector, the p–value becomes $p = 0.00008$. These results lend strong support to the second component of our FORWARD INDUCTION hypothesis; candidates who know that they are facing players who have a choice of which game to play are

more likely to pick the distinct sector in the five-sector game than subjects in the five-sector game without choice.

It is also worth noting that there is a remarkably low number of instances in which one of the non-distinct white sectors was picked in the five-sector game, 2 out of 41 choosers who elected to play the five-sector game and 1 out of 77 candidates. For choosers, a sensible interpretation is that those who were confused about which action to pick in the five-sector game most likely picked the outside option. For candidates, a possible explanation is that there is an aha-effect from contemplating why choosers might want to forego an attractive outside option.

In order to better understand whether choosers were truly motivated by forward-induction considerations we ran another treatment, in which again some players were given a choice between an outside option and playing the five-sector game. Unlike in the case we just discussed though, this treatment was designed to remove forward-induction consideration by not informing conscripts that their partner had a choice, while keeping choosers fully informed about the structure of the game. Therefore we refer to this as the Forward-Induction treatment with *privately* known role assignments. The results from this treatment are summarized in FIGURE 9.

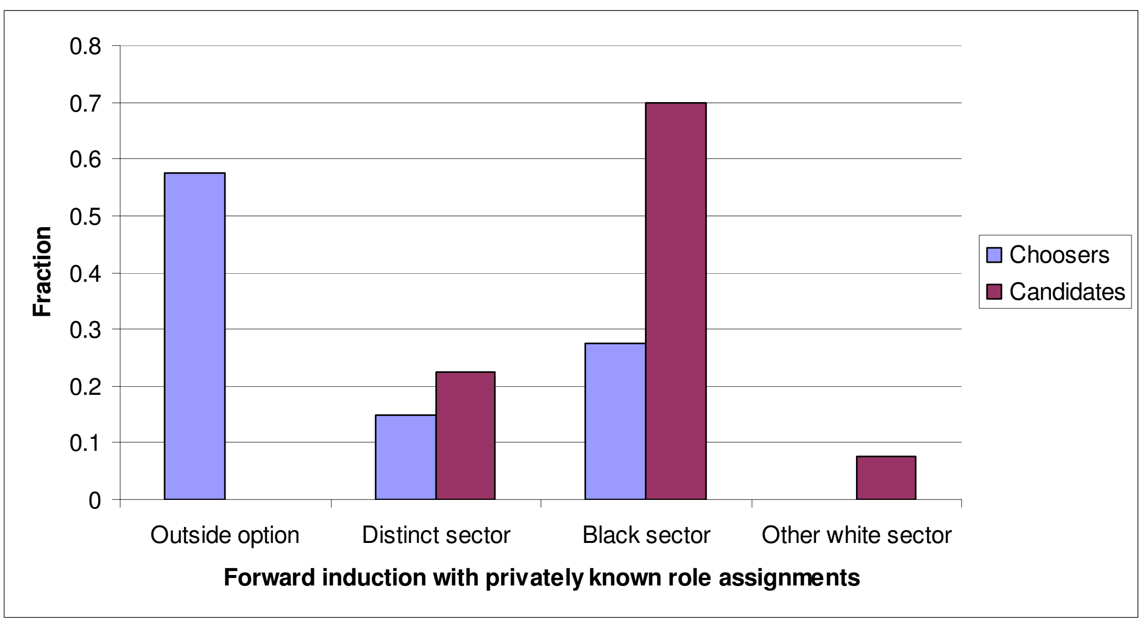


FIGURE 9 Forward Induction - Privately Known Role Assignments

Consistent with the third component of our FORWARD INDUCTION hypothesis, a higher percentage of choosers selected the outside option in the Forward-Induction treatment with privately known role assignments 57.5% (46 out of 80) versus 46.7% (36 out of 77) when role assignments were publicly known, although the difference is not significant at the 10% level, $p = .117$. Conditional on electing to play the five-sector game, choosers were less likely to pick the distinct sector, with 35.3% (12 out of 34) picking the distinct sector when role assignments were private information versus 58.5% (24 of the 41) when role assignments were public information ($p = 0.0377$). This suggests that at least some of the choosers in the forward–induction treatment with publicly known role assignments were guided by forward–induction consideration. At the same time it shows that a substantial number of players elect to play the five–sector game even when there are no signaling opportunities.

5 Related Literature

This paper touches on four areas: Focal points, lack of a common language, lack of common knowledge and forward induction. In this section, we briefly review some of the contributions to these literatures that are relevant for the present paper and were not already discussed.

Focal points were first investigated by Schelling [23], who provides intuition and reports results from some informal experiments. According to Schelling the two prime characteristics of focal points are *conspicuousness* and *uniqueness.* He suggests that finding them may depend more on imagination than on logic. This suggests that Schelling is skeptical of a formal game theoretical investigation of focal points. In particular, he expresses his reservations about the "empirical relevance of mathematical foci." According to Schelling, one should not ascribe to the players in a game the mathematical sophistication of the analyst. For a sophisticated mathematical solution to be focal for a player, that player needs not only to be a mathematician, but must also view his/her playing partners as such.

Schelling's distinction between mathematical and psychological foci is potentially relevant for interpreting the results of our experiment. Essentially, in our games conspicuousness of the black sectors is a psychological phenomenon and competes with the uniqueness of the distinct sector, a mathematical fact. We show that mathematical focal points cannot be dismissed. We demonstrate that the ability to identify mathematical foci varies across the population. Therefore, in situations were agents can self-select, a subpopulation may achieve critical mass; i.e. the mathematical focal point becomes conspicuous for a sufficiently large fraction of the subpopulation for its uniqueness to become decisive.

Sugden [1995] develops a formal theory of focal points by explicitly introducing the labeling of strategies into the analysis. He aims at a "... general theory of how labels can influence decisions in games" (Sugden [1995], p. 534). In a pure coordination game, his theory prescribes that players use decision rules, maps from their private descriptions to a labeled choice, that induce a distribution over choices that maximizes the coordination probability. He argues that in environments with a common culture this prescription often leads to a unique optimal decision rule because of the skewed distribution of the different items mentioned. Our forward–induction results can be interpreted as saying that, when given the opportunity, agents may self-select into common cultures that facilitate coordination.

Mehta, Starmer, and Sugden [1994] examine the concept of a focal point experimentally in pure coordination games. Their objectives are to replicate Schelling's informal experiments, and to discriminate among alternative explanations for coordination success being more frequent than accounted for by pure chance. They distinguish among primary, secondary, and Schelling salience. Primary salience of an action means that (for whatever reason) it is likely to come to mind. An action has secondary salience if it is the optimal reply to one deemed to have primary salience for the playing partner(s). An action has Schelling salience if there is a selection rule that, if used by both players, unambiguously singles out that action as guaranteeing coordination success. They confirm the observation that coordination success is often more frequent than would be suggested by pure chance, and they reject the explanation that this is due to a combination of primary salience and shared cultural experience. They suggest that both secondary and Schelling salience play a role.

We use a variant of VFT, that substitutes symmetry constraints for frames and therefore avoids having to introduce "acts," to organize the data from our experiment. This model is very much in the spirit of and owes an obvious debt to Bacharach's [1993] original formulation. There are three differences, two minor and one more substantive.

The first minor difference is that our treatment of strategies conforms more closely with the conventional way of treating Bayesian games. As in the standard approach, our strategies are functions from private information into actions, conventionally defined. We then let a player's strategy be subject to symmetry constraints. Bacharach instead replaces actions by "acts" that already embody symmetry constraints. We feel that it is an advantage of our approach that we can easily accommodate a richer set of symmetry constraints than those given by partitions.

The second minor difference is in the proposed solution. Our solution is based on recursive optimality. In contrast, Bacharach's solution, translated into our terminology, is a strategy profile that respects the symmetry constraints at every cognition level and for which there is not another strategy profile respecting the same constraints that makes players at one of the levels better off. Such "admissible variable universe equilibria" need not exist. For example, if the fraction of high–cognition players in our circle game is sufficiently large, the high–cognition players strictly prefer the equilibrium in which they pick the distinct sector and the low–cognition players randomize over the three white sectors. Alternatively, one could define an "admissible variable universe equilibrium" as a strategy profile that respects the symmetry constraints at every cognition level and for which there is not another strategy profile respecting the same constraints that makes players at *every* level better off. This will ensure existence but with a high enough probability of high–cognition players in our game will imply multiplicity of equilibria. In examples, Bacharach looks at solutions that in fact satisfy recursive optimality, which suggests to us that we are implementing Bacharach's intent.

The substantive extension of Bacharach's approach it to a permit more general structures on the sets of objects than partitions, or collections of partitions. This allows us to accommodate the circular structure in our games. In other settings, it would help one to incorporate other structures, like relative position, temporal order, size, brightness, compositionality, centrality, adjacency, etc.

Our approach to the players' framing problem is through the language they use to describe a set of objects. Like Crawford and Haller [1990] we model lack of a common language through symmetry constraints on players' strategies. Segal [1999] uses symmetry to model language constraints in a contracting setting. Like Blume [2000] we permit some structure in the language by permitting symmetry restrictions other than complete symmetry. The variant of VFT used in the present paper combines Bacharach's representation of lack of common knowledge with Crawford and Haller [1990] and Blume's [2000] representation of lack of a common language. An alternative way to generalize Bacharach's model would be to use relational structures as in Rubinstein [1996].

Common knowledge of differences in cognition is sometimtes paradoxical; it would not make sense in our setting for low–cognition players to know that they to not recognize the distinct sector as being distinct.[9] This suggests paying closer attention to belief hierarchies, as in Bacharach [1993]. The situation is somewhat analogous to environments in which boundedly rational players have different depths of reasoning, e.g. some playing naive strategies, some playing best responses to naive behavior, some best responding to beliefs that assume best responses to naive play, etc. as in level–k theory.

Forward induction expresses the idea that behavior in a subgame may be influenced by the way in which the subgame was reached. The forward induction idea was introduced by Kohlberg and Mertens [1996]. They show that stable sets of equilibria have the "forward-induction property," i.e. they contain a stable set of any game obtained by eliminating "never-

[9] Of course, it would not be paradoxical for them to know that one of the five sectors is distinct and that others may know which one it is.

weak-best responses." Van Damme [1989] provides an alternative definition of forward induction based on the extensive form of the game. He argues that strategic stability does not fully capture the forward induction logic. Some of the forward induction logic is captured by (iterative) deletion of dominated strategies, e.g. Ben-Porath and Dekel [1992].

The experimental literature has found mixed support for the forward–induction hypothesis. Consistent with the forward induction hypothesis, Cooper, DeJong, Forsythe and Ross (CDFR) [1993] find that providing one player with an outside option in the Battle of the Sexes game predominantly leads to play of that player's favorite equilibrium in the BoS subgame. Somewhat at variance with the forward-induction hypothesis, they find that the outside option in BoS matters even when it yields a payoff lower than the lowest Nash equilibrium in the BoS subgame. Moreover, contrary to a strict interpretation of forward–induction reasoning, CDFR in a coordination game find a large number of choices of the outside option, when FI would suggest that the outside option would never be chosen. Cachon and Camerer [1996] suggest, and provide supporting data, that some observations consistent with backward induction may instead be due to loss avoidance. Furthermore Camerer and Johnson [2004] provide data that suggest that, contrary to what would be implied by forward induction reasoning, responders pay relatively little attention to the choosers' outside–option payoffs. We show that there is a role for forward induction in strategic environments with heterogeneity in cognition and heterogeneity of beliefs about cognition.

6 Conclusion

It appears self-evident that differences in cognition play an important role in many "real-world" strategic interactions. As a consequence, even if interests are perfectly aligned, organizations may not achieve the full benefits of cooperation, and in adversarial settings smarter players may gain persistent advantages.

The contribution of this paper is in showing that participants in our experiments form beliefs about each others' cognition and signal their own cognition if given the opportunity. Failure to coordinate in our strategic environment is shown to result both from failure of cognition and pessimistic beliefs regarding the cognition of others. We extend VFT as the conceptual framework for studying the strategic effect of cognitive differences in pure coordination games and experimentally show that cognitive differences have measurable and predictable effects in these games.

While VFT and cognitive forward–induction do capture some of the central tendencies in our data, the data also poses come challenges for the theory. Among the more important ones are:

Why does it seem that prior play against a partner worsens performance when playing against oneself? Why is there no statistically significant difference between the proportion of distinct–sector choices in the Partner–Adjacent and Self–Adjacent treatments? Why is the difference in proportions of choosers selecting the outside option between the cases of publicly and privately known role–assignments not statistically significant? While it may suffice to simply have more data to answer the latter two questions, the first puzzle may require us to look beyond the confines of VFT.

In future work along these lines it would also be interesting to investigate how agents, or groups of agents, can improve their strategic sophistication; whether strategic sophistication in one domain carries over to other domains; and, how agents deal with similar cognitive tasks in conflict situations.

References

Bacharach, M. [1993], "Variable Universe Games," in K. Binmore, A. Kirman, and P. Tani (eds.), *Frontiers of Game Theory,* Cambridge, MA: MIT Press.

Bacharach, M. and M. Bernasconi [1997] "The Variable Frame Theory of Focal Points: An Experimental Study," *Games and Economic Behavior,* **19**, 1-45.

Bacharach M. and D.O. Stahl [2000] "Variable-Frame Level-n Theory," *Games and Economic Behavior,* **32**, 220-246.

Ben-Porath, E. and E. Dekel [1992], "Signaling Future Actions and the Potential for Self-Sacrifice," *Journal of Economic Theory,* **57**, 36-51.

Board, O. and K.–S. Chung [2007], "Object–Based Unawareness," University of Pittsburgh working paper.

Blume, A. [2000], "Coordination and Learning with a Partial Language," *Journal of Economic Theory,* **95**, 1-36.

Blume, A. and U. Gneezy [2000a], "An Experimental Investigation of Optimal Learning in Coordination Games," *Journal of Economic Theory,* **90**, 161-172.

Blume, A. and U. Gneezy [2000b], "Cognitive Forward Induction and Coordination without Common Knowledge: Theory and Evidence," University of Pittsburgh and Technion Working Paper.

Blume, A., D. DeJong, Y.-G. Kim, and G. Sprinkle [2001], "Evolution of Communication with Partial Common Interest," *Games and Economic Behavior, Games and Economic Behavior*, **37**, Issue 1, October 2001, 79-120.

Cachon, G.P. and C.F. Camerer [1996], "Loss-Avoidance and Forward Induction in Experimental Coordination Games," *The Quarterly Journal of Economics,* **111**, 165-194.

Camerer, C., and E. Johnson [2004], "Thinking Backward and Forward in Games," in I. Brocas and J. Castillo (Eds.), *The Psychology of Economic Decisions*, Oxford: Oxford University Press.

Casajus, A. [2001] *Focal Points in Framed Games, Lecture Notes in Economics and Mathematical Systems 499*, Berlin: Springer-Verlag.

Cooper, R., D.V. DeJong, R. Forsythe and T.W. Ross [1993], "Alternative Institutions for Resolving Coordination Problems: Experimental Evidence on Forward Induction and Preplay Communication," in J.W. Friedman (editor), *Problems of Coordination in Economic Activity,* Boston, Dordrecht, London: Kluwer Economic Publishers.

Crawford, V. [2003], "Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions," *The American Economic Review*, **93**, 133-149.

Crawford, V. and H. Haller [1990], "Learning how to Cooperate: Optimal Play in Repeated Coordination Games," *Econometrica,* **58**, 581-596.

Dekel, E., B.L. Lipman and A. Rustichini [1998], "Standard State-Space Models Preclude Unawareness," *Econometrica,* **66**, 159-173.

Fagin R. and J.Y. Halpern [1988], "Belief, Awareness and Limited Reasoning," *Artificial Intelligence,* **34**, 39-76.

Feinberg, Y. [2005], "Games with Incomplete Awareness," Stanford University working paper.

Geanakoplos, J. [1989], "Game Theory without Partitions, and Applications to Speculation and Consensus," Yale University Working Paper.

Halpern, J.Y. and L.C. Rêgo [2005], "Interactive awareness revisited," *Proceedings of Tenth Conference on Theoretical Aspects of Rationality and Knowledge,* 78-91.

Heifetz, A., M. Meier and B. Schipper [2006], "Interactive Unawareness," *Journal of Economic Theory*, **130**, 78-94.

Heifetz, A., M. Meier and B. Schipper [2007], "Unawareness, Beliefs and Games," University of California–Davis working paper.

Kohlberg, E. and J.-F. Mertens [1986], "On the Strategic Stability of Equilibria," *Econometrica,* **54**, 1003-1037.

Lipman, B.L. [1999], "Decision Theory without Logical Omniscience: Toward an Axiomatic Framework for Bounded Rationality," *The Review of Economic Studies*, **66**, 339–361.

Mehta, J., C. Starmer, and R. Sugden [1994], "The Nature of Salience: An Experimental Investigation of Pure Coordination Games," *American Economic Review,* **84**, 658-673.

Modica, S. and A. Rustichini [1999], “Unawareness and Partitional Information Structures,” *Games and Economic Behavior,* **27** 265-298.

Nagel, R. [1995], “Unravelling in Guessing Games: An Experimental Study,” *American Economic Review,* **85**, 1313-1326.

Rubinstein, A. [1996], “Why are Certain Properties of Binary Relations Relatively More Common in Natural Language?” *Econometrica,* **64**, 343-355.

Schelling, T. [1960], *The Strategy of Conflict,* Cambridge, MA: Harvard University Press.

Segal, I. [1999], “Complexity and Renegotiation: A Foundation for Incomplete Contracts,” *Review of Economic Studies,* **66**, 57-82.

Stahl, D. and P. Wilson [1994], “Experimental Evidence on Players’ Models of Other Players,” *Journal of Economic Behavior and Organization*, **25**, 309-327.

Stahl, D.O. and P.W. Wilson [1995] “On Players’ Models of Other Players: Theory and Experimental Evidence,” *Games and Economic Behavior,* **10**, 218-254.

Sugden, R. [1995], “A Theory of Focal Points,” *The Economic Journal,* **105**, 533-550.

van Damme, E. [1989], “Stable Equilibria and Forward Induction,” *Journal of Economic Theory,* **48**, 476-496.

Appendix 1: Raw Data

Table II

Partner - Separated

Observation	Stage 1	Stage 2	Observation	Stage 1	Stage 2
1	d	d	16	b	b
2	d	d	17	b	b
3	d	d	18	b	b
4	d	d	19	b	b
5	d	d	20	b	b
6	d	d	21	b	b
7	d	d	22	b	b
8	d	d	23	b	w
9	d	b	24	w	w
10	d	b	25	w	w
11	d	b	26	w	w
12	b	d	27	w	b
13	b	d	28	w	b
14	b	b	29	w	b
15	b	b	30	w	b

Table III

Self - Separated

Obs.	Stage 1	Stage 2	With Partner	Obs.	Stage 1	Stage 2	With Partner
1	d	d	d	16	d	d	d
2	d	d	d	17	d	d	d
3	d	d	d	18	w	w	d
4	d	d	d	19	w	w	d
5	d	d	d	20	w	w	d
6	d	d	d	21	w	w	d
7	d	d	d	22	b	b	d
8	d	d	d	23	w	b	d
9	d	d	d	24	w	d	d
10	d	d	d	25	w	w	w
11	d	d	d	26	w	w	w
12	d	d	d	27	b	b	w
13	d	d	d	28	b	b	b
14	d	d	d	29	b	b	b
15	d	d	d	30	d	d	b

Table IV

Partner–Separated followed by Self–Separated

Obs.	Stage 1 Partner	Stage 2 Partner	Stage 3 Self	Stage 4 Self	Obs.	Stage 1 Partner	Stage 2 Partner	Stage 3 Self	Stage 4 Self
1	d	d	d	d	16	b	b	b	b
2	d	d	d	d	17	b	b	b	b
3	d	d	d	d	18	b	b	b	b
4	d	d	d	d	19	b	b	b	b
5	d	d	d	d	20	b	b	b	b
6	d	d	d	d	21	b	b	b	b
7	d	d	b	b	22	b	b	b	b
8	d	b	d	d	23	b	b	w	w
9	d	b	d	d	24	b	b	w	b
10	d	b	d	d	25	w	b	b	b
11	d	b	d	b	26	w	d	d	d
12	d	b	d	b	27	w	w	b	b
13	b	b	d	d	28	w	w	w	w
14	b	b	d	d	29	w	w	w	w
15	b	b	d	d	30	w	w	d	d

Table V

Partner - Adjacent

Obs.	Stage 1	Stage 2	Stage 3	Obs.	Stage 1	Stage 2	Stage 3
1	d	d	d	16	b	b	b
2	d	d	d	17	b	b	d
3	d	b	b	18	b	b	w
4	b	d	b	19	b	b	w
5	b	b	b	20	b	w	w
6	b	b	b	21	b	w	b
7	b	b	b	22	w	b	b
8	b	b	b	23	w	b	b
9	b	b	b	24	d	b	w
10	b	b	b	25	w	b	w
11	b	b	b	26	w	w	w
12	b	b	b	27	w	w	w
13	b	b	b	28	w	w	w
14	b	b	b	29	w	w	w
15	b	b	b	30	w	d	w

Table VI

Self - Adjacent

Obs.	Stage 1	Stage 2	With Partner	Obs.	Stage 1	Stage 2	With Partner
1	d	d	d	16	b	b	b
2	d	d	d	17	b	b	b
3	d	d	d	18	b	b	b
4	d	d	d	19	b	b	b
5	d	d	d	20	b	b	d
6	d	d	d	21	b	b	w
7	d	d	d	22	b	b	w
8	d	d	b	23	w	w	w
9	b	b	b	24	w	w	w
10	b	b	b	25	w	w	w
11	b	b	b	26	w	w	w
12	b	b	b	27	w	w	w
13	b	b	b	28	w	w	b
14	b	b	b	29	w	w	d
15	b	b	b	30	w	w	d

Table VII

Forward induction when all participants have a choice

Outside option	Distinct sector	Black sector	Other white sector	Total
66	27	5	2	100

Table VIII

Forward induction with publicly known role assignments

	Outside option	Distinct sector	Black sector	Other white sector	Total
Choosers	36	24	15	2	77
Candidates	n/a	45	31	1	77

Table IX

Forward induction with privately known role assignments

	Outside option	Distinct sector	Black sector	Other white sector	Total
Choosers	46	12	22	0	80
Candidates	n/a	18	56	6	80

Appendix 2

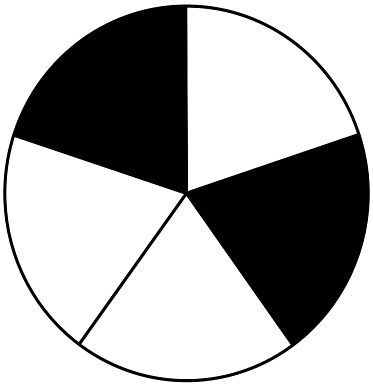
Instructions for

1. Play against a partner (Translation from Hebrew)
2. Play against oneself (Translation from Hebrew)
3. Forward induction where all participants have a choice (Translation from Hebrew)
4. Forward induction with publicly known role assignments
5. Forward induction with privately known role assignments

Instructions

Welcome to this experiment in decision making. Soon you will be randomly matched with another student. The experiment will take about 20 minutes and you may earn up to 40 Shekel in it.

In the first stage of the experiment one person in each couple will be asked to make a simple decision: to choose one out of the five sectors of a pie chart similar to this.



The actual pie will be a two-sided pie, in which the two sides are identical {At this point the subjects were given the actual pie chart used in the experiment in order to observe it}.

We will mark this choice on the inside of the chart. Then we will flip the chart randomly, and show it to the other subject. The second subject will not know whether (s)he observes the chart from the same side as the first subject or not. (S)he will also not see the choice of the first player. We will then ask the second subject also to choose one sector out of the five.

Payment: Each of the two subjects will be paid 20 shekels if they will choose the same sector and nothing otherwise.

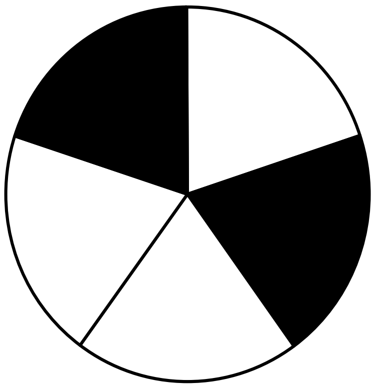
We will then repeat this experiment once more in exactly the same way. We will not let you know the outcome of the first experiment before choosing in the second. At the end of the second experiment the payment for the two experiments will be calculated and you will be paid privately and in cash.

Do you have any questions?

Instructions

Welcome to this experiment in decision making. The experiment will take about 20 minutes and you may earn up to 60 Shekel in it.

In the first stage of the experiment you will be asked to make a simple decision: to choose one out of the five sectors of a pie chart similar to this.



The actual pie will be a two-sided pie, in which the two sides are identical. {At this point the subjects were given the actual pie chart used in the experiment in order to observe it}.

We will mark this choice on the inside of the chart. Then we will flip the chart randomly, and show it to you again. You will not know whether you observe the chart from the same side as before or not. We will then ask you to choose again one sector out of the five.

Payment: You will be paid 20 Shekel if you choose the same sector in the two stages and nothing otherwise.

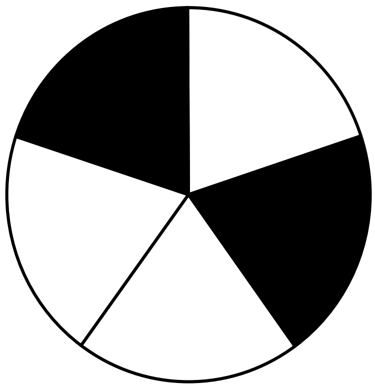
We will then give you the instructions for the second part of the experiment. At the end of the second experiment the payment for the two experiments will be calculated and you will be paid privately and in cash.

Do you have any questions?

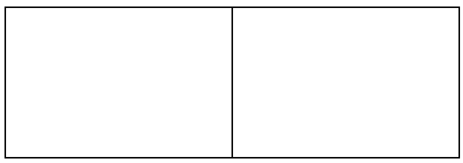
Instructions

Welcome to this experiment in decision making. The experiment will take about 20 minutes and you may earn up to 40 Shekel in it.

We will first ask you to choose one of two objects, similar to the following:



Or



You will be matched with another subject who chose the same object as you did.

If you choose the pie, then one person in each couple will be asked to make a simple decision: to choose one out of the five sectors of a pie chart similar to this. The actual pie will be a two-sided pie, in which the two sides are identical. {At this point the subjects were given the actual pie chart used in the experiment in order to observe it}.

We will mark this choice on the inside of the chart. Then we will flip the chart randomly, and show it to the other subject. The second subject will not know whether (s)he observes the chart from the same side as the first subject or not. (S)he will also not see the choice of the first player. We will then ask the second subject also to choose one sector out of the five.

<u>If you choose the rectangle</u>, then the experiment will be done in a similar way but with the rectangle.

Payment: Regardless of the object chosen, each of the two subjects will be paid 40 Shekel if they choose the same sector and nothing otherwise.

Remark: In case we will have an uneven number of subjects choosing the pie, one of them will be chosen randomly to participate in the rectangle experiment.

At the end of the experiment the payment will be calculated and you will be paid privately and in cash.

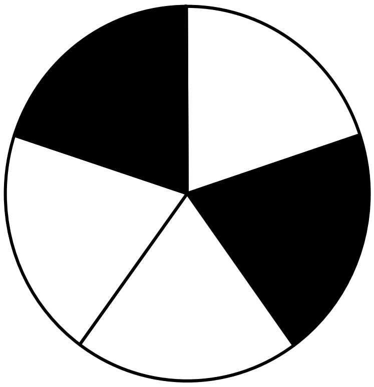
Do you have any questions?

Instructions for Student A

Welcome to this experiment in decision making. Soon you will be randomly matched with another student. You are called Student A and the other student is called Student B. The experiment will take about 15 minutes and you may earn up to $10 in it.

In the first stage of the experiment you will be asked to make a simple decision between the following two options:

Option 1: choose one out of the five sectors of a pie chart similar to this.



The actual pie will be a two-sided pie, in which the two sides are identical.

We will mark this choice on the inside of the chart. Then we will flip the chart randomly, and show it to Student B. Student B will not know whether (s)he observes the chart from the same side as you or not. (S)he will also not see your choice. We will then ask Student B also to choose one sector out of the five.

Payment: Each of you will be paid $10 if you choose the same sector and nothing otherwise.

Or, you may choose option 2.

Option 2: be paid $6 for participating in the experiment

If you choose option 1, than your payment will depend on both your choices as described above. If you choose option 2 then each of you will be paid $6.

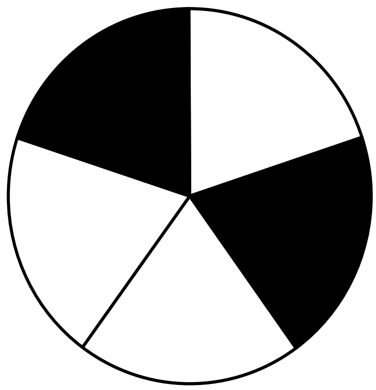
Do you have any questions?

Instructions for Student B

Welcome to this experiment in decision making. Soon you will be randomly matched with another student. You are called Student B and the other student is called Student A. The experiment will take about 15 minutes and you may earn up to $10 in it.

In the first stage of the experiment we ask Student A to make a simple decision between the following two options:

<ins>Option 1</ins>: choose one out of the five sectors of a pie chart similar to this.



The actual pie will be a two-sided pie, in which the two sides are identical.

We will mark Student A's choice on the inside of the chart. Then we will flip the chart randomly, and show it to you. You will not know whether (s)he observes the chart from the same side as you or not. You will also not see Student A's choice. We will then ask you also to choose one sector out of the five.

Payment: Each of you will be paid $10 if you choose the same sector and nothing otherwise.

Or, Student A may choose option 2.

Option 2: be paid $6 for participating in the experiment

If Student A chooses option 1, than your payment will depend on both your choices as described above. If Student A chooses option 2 then each of you will be paid $6.

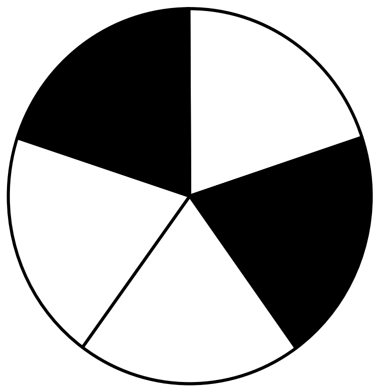
Do you have any questions?

<u>Instructions for Student A</u>

Welcome to this experiment in decision making. Soon you will be randomly matched with another student. You are called Student A and the other student is called Student B. The experiment will take about 15 minutes and you may earn up to $10 in it.

In the first stage of the experiment you will be asked to make a simple decision between the following two options:

<u>Option 1</u>: choose one out of the five sectors of a pie chart similar to this.



The actual pie will be a two-sided pie, in which the two sides are identical.

We will mark this choice on the inside of the chart. Then we will flip the chart randomly, and show it to Student B. Student B will not know whether (s)he observes the chart from the same side as you or not. (S)he will also not see your choice. We will then ask Student B also to choose one sector out of the five.

Payment: Each of you will be paid $10 if you choose the same sector and nothing otherwise.

Or, you may choose option 2.

Option 2: be paid $6 for participating in the experiment

If you choose option 1, than your payment will depend on both your choices as described above. If you choose option 2 then each of you will be paid $6.

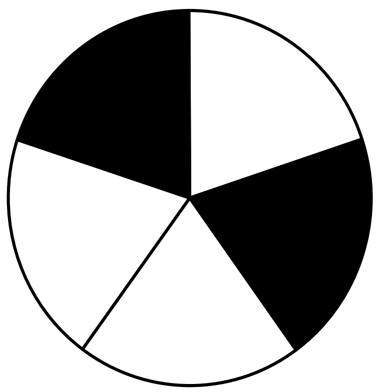
In either case Student B will not know that you had a choice between options 1 and 2.

Do you have any questions?

<u>Instructions for Student B</u>

Welcome to this experiment in decision making. Soon you will be randomly matched with another student. You are called Student B and the other student is called Student A. The experiment will take about 15 minutes and you may earn up to $10 in it.

In the first stage of the experiment we ask Student A to choose one out of the five sectors of a pie chart similar to this.



The actual pie will be a two-sided pie, in which the two sides are identical.

We will mark Student A's choice on the inside of the chart. Then we will flip the chart randomly, and show it to you. You will not know whether (s)he observes the chart from the same side as you or not. You will also not see Student A's choice. We will then ask you also to choose one sector out of the five.

Payment: Each of you will be paid $10 if you choose the same sector and nothing otherwise.

Do you have any questions?