Learning Categorizations of Strategy Spaces

Vessela Daskalova† and Nicolaas J. Vriend‡

25 February 2014

Abstract

We consider categorization of the strategy space in games, in which some attributes of the game can serve as a basis for distinguishing the available strategies. There may be many different ways to partition the strategy space in a given game and some of these ways may be nested in one another. We present a model in which agents learn how to perceive the strategic situation they are facing. Although this general model can be applied to a wider class of games, we focus on coordination games. We show which categorizations of the strategy space help to solve the equilibrium selection problem most efficiently and how agents may learn them through reinforcement. We also fit the model to data from two laboratory experiments and find that it accounts for their empirical findings.

Keywords: Strategy Space, Frames, Categories, Variable Frame Theory, Reinforcement Learning, Coordination Games, Equilibrium Selection

JEL Classification: C72, C63, C91

1 Introduction

Both casual introspection and experimental evidence from psychology show that categorization is a widespread phenomenon. One important area in which people may categorize is games. Players may categorize other players (Azrieli, 2009), they may distinguish various categories of games (Mengel, 2012) or they may bundle the options of play of other players at a given node into categories (Jehiel, 2005).

†Cambridge-INET, University of Cambridge
‡Queen Mary, University of London
In this paper we consider categorization of the strategy space in games. That is, a player groups her strategies into categories. Categorization implies that the person treats objects that she has placed in the same category in the same way. Here the objects are the strategies.

Categorization of the strategy space seems relevant in many situations of strategic interaction. For example, managers may categorize the strategies available to them in market interactions, political leaders may categorize the strategies available to them in domestic and foreign policy, generals may categorize the strategies available to them in war, and players and coaches may categorize the strategies available to them in sports. Typically a number of different categorizations of the strategy space are possible. Depending on the task at hand some categorizations may be better than other categorizations. And in many games which categorizations are best depends on the categorizations chosen by the other players.

The main goal of this paper is to improve our understanding of how people decide which categorizations of a strategy space to use. We present a model of agents choosing among alternative categorizations of a given strategy space. And we confront this model with some empirical data, examining how the model may help us understand behaviour in two laboratory experiments (Bosch-Domènech and Vriend (2013) (BDV) and Blume and Gneezy (2010) (BG)). We also analyze the dynamic properties of our model, e.g. its long run behavior and its capacity to achieve maximum efficiency.

Our model is a dynamic model based on reinforcement learning. That is, players learn which categorizations and categories of strategies to use based on their past experience, and they are more likely to choose those that have performed better in the past. We study how views of the world that are seen as useful, as well as how the corresponding outcomes, emerge in the process of social interaction. Taking such a dynamic perspective has several advantages. First, using reinforcement learning allows us to keep the model modest in terms of cognitive assumptions. One can thus expect that a relatively broad class of more complex cognitive models may share some of the behavioural properties of our model. Second, we focus on coordination games, i.e. games with a multiplicity of equilibria. This multiplicity of equilibria in coordination games will be apparent also when it comes to categorizations. Instead of aiming for equilibrium refinements, a study of some plausible learning dynamics may help to shed some light on the outcomes one may expect in such games. Third, as we will see in our analysis, learning dynamics may present some additional reasons to favour one categorization over another, reasons that may not be obvious in a static analysis.

The remainder of this paper is organized as follows. In section 2 we discuss
the paper’s relation to the literature. In section 3 we introduce the BDV and BG games and we discuss the different possible categorizations of the strategy space in these games, as well as the NE in expected payoffs that arise under the set of possible representations of the strategy space. Section 4 presents our learning model. In section we analyze the model in relation to the BDV experiment, and we show that the model can be applied to the BG game, as well. Section S6 concludes.

2 Relation to the Literature

As mentioned in the introduction, Azrieli (2009), Mengel (2012), and Jehiel (2005) also consider categorization in games. Unlike those papers we focus on categorization of the strategy space. Our work is most related to the literature on equilibrium selection in coordination games, in particular to Variable Frame Theory (Bacharach, 1993; Bacharach and Bernasconi, 1997; Bacharach and Stahl, 2000), to Janssen (2001), and to Sugden (1995). We first discuss what the paper has in common with these papers and we then explain in what ways it is complementary to them.

The underlying idea of the above papers (Bacharach, 1993; Bacharach and Bernasconi, 1997; Bacharach and Stahl, 2000; Janssen, 2001; Sugden, 1995) is that the standard normal form representation of the game that is usually used to depict a strategic situation does not necessarily capture the way a player thinks of the situation. In many cases, there are alternative ways in which a player may describe the strategy space to herself. The player’s description of her options depends on which attributes a player is able to perceive (in Variable Frame Theory and in Janssen (2001)) or it depends on which labels she uses to describe the strategies to herself (Sugden, 1995). A key insight of this literature is that the use of these alternative descriptions of the game may provide the basis for selecting among multiple equilibria.

To illustrate, consider the following example from VFT. Two players play a matching game. In this game a number of objects are given and each of the two players has to pick an object independently from the other. If both players pick the same object they both get the same positive payoff, and if they pick different objects they both get a payoff of zero. Let us assume that the set that they have to choose from consists of five objects and that coordinating on an object brings the same payoff to both players regardless of which option they coordinate on. Thus, under the standard normal form representation of the game there are five payoff equivalent NE and players face an equilibrium selection problem. If the payoff in each of the possible equilibria is 1 for each player and if each of the possible equilibria is equally likely to be selected, then the expected payoff in this game is 0.2 for each player.
VFT focuses on games in which players can distinguish some attributes of the objects. For example, a player may view the objects through a colour frame, meaning that she splits the objects into different classes depending on their color. In the above example if one object is blue and the rest are white, then the players may distinguish between the option “pick the blue object” and the alternative “pick any white object”. If both players view the game under the color frame and this is common knowledge, then the choice profile (“pick the blue object”, “pick the blue object”) entails the highest expected payoffs for the players. VFT predicts that this is the solution to the game. There are two principles that underly this solution. The first one is what Bacharach and Bernasconi (1997) call Symmetry Disqualification, and Janssen (2001) refers to as the Principle of Insufficient Reason.\(^1\) Symmetry Disqualification says that if a player sees insufficient reason to distinguish between certain objects, then she will treat them in the same way. Thus, she will randomize uniform randomly among them. Symmetry Disqualification is also related to the concept of attainable strategies in Crawford and Haller (1990). An attainable strategy is a strategy such that a player plays the actions that she does not distinguish among with equal probability. The second principle invoked by VFT is the Principle of Payoff Dominance. It dictates that players will select the Pareto-superior equilibrium in expected payoffs among all equilibria that they are aware of.

Janssen (2001) offers an extension of VFT related to the question which objects players treat in the same way. In VFT players treat objects that are description symmetric (in the above example, the white objects) according to the Principle of Insufficient Reason. In Janssen (2001) not only objects that are description symmetric but also objects that are payoff symmetric are treated according to the Principle of Insufficient Reason. We reproduce an example from Janssen (2001)’s paper to illustrate. Consider a matching game in which there are two red, one blue, one green, and one yellow object. According to VFT there are four NE in expected payoffs: both players choosing “pick any red”, both choosing “pick the blue”, both choosing “pick the green”, both choosing “pick the yellow”. The latter three equilibria are equivalent in terms of efficiency and Pareto-superior to both choosing “pick any red”. VFT does not offer a solution to the equilibrium selection problem players face in selecting one of these three equilibria. Janssen (2001) argues that although they are not description symmetric, the three options (“pick the blue”, “pick the green”, “pick the yellow”) are payoff symmetric, thus a player has no reason to treat them differently and she should choose among them according to the Principle of Insufficient Reason. If she treats the three equilibria according to the Principle of Insufficient Reason, then (“pick any red, pick any

---

\(^1\)An early discussion of the Principle of Insufficient Reason is attributed to Jakob Bernoulli. Later discussions include Keynes (2007).
red”) is Pareto-superior in expected payoffs and thus a solution to the game.

In Sugden (1995) each player can always distinguish her options by assigning private labels to them. Thus, a player does not use the Principle of Insufficient Reason to select among her options. Whether coordination between players is successful or not depends on the extent to which players’ private labels are common knowledge. This idea will play a role also in the Blume and Gneezy (2010) game that we discuss later.

This paper is closely related to the literature described above. We also assume that there may be many alternative representations of the strategy space of a given game. We use the concept of a categorization, i.e. a set of categories of the strategy space, that is very similar to the frame concept in VFT. As in Janssen (2001) we assume that players may form categories of strategies both on the basis of description and on the basis of payoff symmetry. Note that treating objects that are similar to each other according to the Principle of Insufficient Reason is equivalent to putting them in the same category.

There are two main differences between this work and the above literature. First, in our model we focus on situations in which some attributes are distinguished relatively naturally. Given these attributes, there is a collection of possible representations of the situation, i.e. of different categorizations of the strategy space. In contrast to the literature, we assume that ex ante all these naturally arising categorizations are available to each player. This differs from VFT and from Janssen (2001), in which a player has one possible representation of the situation based on the families of attributes she has the cognitive capacity of distinguishing. Thus, our model is not one of limited cognitive capacities. Instead the key question on which we focus is how agents learn to make active use of some views of the world while leaving others unused.

The second main difference is that our model is dynamic. To deal with the issue of selecting among a multiplicity of equilibria, the approach in most of the literature has been to rely on principles such as the Principle of Coordination (VFT), Principle of Collective Rationality (Sugden, 1995) or Principle of Individual Team Member Rationality (Janssen, 2001). In contrast to such a static approach we follow a dynamic approach based on reinforcement learning (see e.g. Roth and Erev (1995); Erev and Roth (1998)) with players being more likely to select categorizations that have been more successful in the past. Thus, we complement the static analyses in the literature by looking at the question which categorization the players are likely to choose and which outcomes are likely to emerge in the process of social interaction. Our dynamic approach means that we do not have to

---

2Note that we focus on learning categorizations of the strategy space that are useful for coordination in a series of one-off interactions. Learning to coordinate in repeated interactions has been considered by Crawford and Haller (1990) and by Goyal and Janssen (1996) among others.
rely on a principle such as the Principle of Coordination (VFT), the Principle of Individual Team Member Rationality (Janssen, 2001) or the Principle of Collective Rationality (Sugden, 1995) to select among a multiplicity of equilibria. Rather than postulating that players coordinate on the Pareto superior equilibrium or that they reason as a team, we study whether and in what cases the Pareto superior equilibrium emerges as a result of the interaction. By looking at which equilibria emerge in the process of social interaction we are also able to learn more about the advantages and disadvantages of different categorizations when it comes specifically to learning.

A further minor difference with the literature is that in some of these models players with limited cognitive capacities have beliefs about the (weakly) lower cognitive capacities of their opponents (see e.g. Janssen (2001) or Bacharach and Stahl (2000)). In our model players have no explicit beliefs about which categorizations the other players use.3

3 Model

In this section we describe the general set-up of our model. The starting point is that each game has a strategy space, i.e. a set of possible options that players choose from. We assume that players categorize the strategies available to them. That is, a player groups her strategies into categories. A categorization or frame is a set of categories of strategies and represents a player’s view of the strategy space. Depending on which attributes of the strategy space a player pays attention to, she may use different categorizations of the strategy space, i.e. different views of the world.4

Note that sometimes in the cases that we analyze later distinguishing one attribute may be necessary before another attribute can be distinguished. Thus, some categories within a categorization may be distinguished only after distinguishing a coarser category that they all belong to. For example, before distinguishing between “attacking the right flank” or “attacking the left flank”, football players have to distinguish the category of strategies “attacking the flanks”.

We assume that ex ante players can observe all possible attributes of the situation they are facing. That is, any attribute that we as modelers distinguish will be distinguished by the players.5 The central question of our model is,

3One could argue that by basing their decisions on their past experience players form implicit beliefs about the potential usefulness of different categorizations and categories.
4We use the terms frame, categorization, and view of the world interchangeably.
5We assume that players recognize all possible attributes. Modeling the learning of the concepts as such is beyond the scope of this paper.
assuming that ex ante players can distinguish all possible attributes of a situation, which attributes do they find worth paying attention to? That is, which of the many possible categorizations of the strategy space or which views of the world do they find most useful?

In the model players play a series of one-shot games (same game each time). Each period all players are randomly and anonymously matched in pairs. A player has a collection of possible categorizations of the strategy space that are derived from the structure of the game. Each categorization is a set of categories. Each period each player needs to choose a categorization to use. She needs to choose a category from the categorization she has selected. And she needs to choose a strategy from the category chosen. Every categorization and category have a strength that is determined by their past performance. Players choose which categorizations and which categories to use each period on the basis of the perceived strengths of the available options. Thus, we need to specify first, how these perceived strengths are determined, and second, how they will be taken into account in the choices. We will start with the latter.

In each period an agent first chooses which categorization to use with the logit rule (McFadden, 1974; Cramer, 2003). The probability of a categorization \( i \) to be chosen in period \( t \) is equal to

\[
p_t(i) = \frac{e^{\beta s_t(i)}}{\sum_j e^{\beta s_t(j)}},
\]

where \( s_t(i) \) (\( s_t(j) \)) is the strength of categorization \( i \) (\( j \)) in period \( t \) and \( \beta \) is a parameter determining the sensitivity of choice to strengths. For \( \beta = 0 \) choice is uniform random, all categorizations are chosen with equal probability. We use \( \beta > 0 \) which means that categorizations that had a better performance in the past are more likely to be chosen. After choosing a categorization to use a player also chooses a category from this categorization. This choice is made again with the logit rule (same formula as above but \( s_t(i) \) denotes category rather than categorization strength). A player then chooses a strategy within the category chosen according to the Principle of Insufficient Reason. That is, players choose uniform randomly among the strategies they have put in the category. The perceived strengths of the categories and categorizations are determined as follows. The initial strengths of categories and categorizations are captured by the parameter \( s_0 \). The strength of a category or categorization in period \( t \), \( s_t \) is updated whenever the category or categorization is used according to the following rule:

\[
s_t(i) = s_{t-1}(i)(1 - \alpha) + \pi_t \alpha,
\]

\( ^6 \)As will be explained below, strategies within a category will be chosen uniform randomly.
where $s_{t-1}(i)$ is the strength of this categorization (category) in the previous period, $\alpha$ is the weight the agent places on the latest interaction, and $\pi_t$ is the payoff from the interaction in period $t$. The higher the $\alpha$, the greater the weight the agent places on more recent experiences. Note that some categories may exist in more than one categorization of the agent. The strengths of a category is updated in all categorizations under which it exists.

To sum up, there are three free parameters in this model: the initial strength $s_0$, the learning rate $\alpha$, and the temperature in the logit rule $\beta$. These parameters appear both at the categorization and at the category level. In principle the parameter values for $s_0$, for $\alpha$, and for $\beta$ could be different for the categorization choice and for the category choice. Furthermore, there could be a specific initial strength for each category and categorization. Parameter values could also differ between individual players and they could vary over time. For the sake of parsimony we decided to use for each instance in our model only one parameter value for all initial strengths $s_0$, only one parameter value for all updating strengths $\alpha$, and only one parameter value for the choice parameter $\beta$.

Note that the agents do not receive any feedback on the categorizations other players use, on the categories or on the strategies chosen by them. Neither do they receive feedback on the other player’s payoff. Also, they do not form any explicit beliefs about the categorizations and categories that other players use.

4 Two Coordination Games

In this section we present two games used in laboratory experiments with human subjects, that we investigate with our model in section 5. They are interesting for several reasons. First, attributes and hence categories of strategies in these games can be distinguished quite naturally. Second, there is a hierarchy of attributes. That is, some attributes can only be distinguished after others have been distinguished. Hence sets of possible categorizations present themselves naturally. Third, applying our model to experimental data on the behavior of human subjects allows us to gain some insights into the categorizations of the strategy space that are consistent with the dynamics of behavior observed in the experiment.

The two games come from the experiments by Bosch-Domènech and Vriend (2013) and by Blume and Gneezy (2010). Both games are coordination games, i.e. games in which players face an equilibrium selection problem. Moreover, the equilibrium selection problem is a difficult one, as all equilibria are equally efficient.
4.1 Game of Bosch-Domènech and Vriend (2013)

We first introduce the game played in the BDV experiment. At the beginning of the experiment players are randomly assigned to be either a row player or a column player. Each player receives a payoff matrix of the game. Players are randomly and anonymously matched in pairs each period. A player’s task is to independently choose one of 15 possible rows/columns of the payoff matrix. The payoff matrix is presented in Figure 1a, where the payoff in each cell is the payoff that is given to each player. There are overall 30 payoff equivalent NE. These NE are scattered in the payoff matrix in a way such that there are always two NE per row/column in an attempt to avoid any equilibrium being more conspicuous than others (e.g. there are no NE in the corners of the payoff matrix). The matrix shown in Figure 1a is the version of the game that was just described and that is used in the control treatment.

There are three other treatments. In each of them a focal point is introduced by shaving the payoff of one of the NE as in Figure 1b. Note that in Figure 1b one of the cells contains a payoff of 87. This outcome is focal, that is, it is unique and conspicuous. After shaving the NE payoff, the corresponding action profile is no longer a NE. Given that one player plays the focal action, the other player has an incentive to deviate to what BDV call an Associated NE (ANE). That is, the NE in the same row/column as the focal payoff. Note, however, that there are two ANE, one in the same row and one in the same column as the focal point. If both players deviate from the focal action in the hope of realizing an ANE, they would miscoordinate. Thus, in the the focal treatments instead of 30 NE, there are 29 NE and one non-equilibrium focal point. This focal point takes the values 46, 87, and 99 in the different treatments, respectively. The payoff matrices for the 46 and 99 treatments are identical to the payoff matrix shown in Figure 1b apart from the magnitude of the focal payoff being 46 or 99 rather than 87.
We now discuss the different possible frames or categorizations of the strategy space in the BDV game with a focal point.\footnote{A categorization in our model is a hierarchical set of categories. However, note that our hierarchies are not to be confused with the hierarchies discussed in the literature on level-k reasoning (Nagel, 1995; Stahl and Wilson, 1994, 1995; Camerer et al., 2004; Costa-Gomes and Crawford, 2006). In the literature on level-k reasoning, players differ in their depth of reasoning in a strategic situation. Each higher level of reasoning is a best response to players’ using the previous level of reasoning. In some versions of level-k players may differ also in the beliefs that they hold about the distribution of the levels of reasoning in the population. Note that in our case a higher frame is not a best response to lower frames. In our model frames are based on attributes of the game that players perceive.} They are illustrated in Figure 2. The coarsest possible way to categorize the strategy space is to put all strategies in one category, the category “any”. We call this categorization F0. Putting all strategies in one category implies that a player chooses uniform randomly among them. However, a player may have a finer view of the strategy space. She may observe that there is a payoff in one of the cells of the payoff matrix that differs from all others, i.e. she may note the focal payoff. This means that she distinguishes the “focal” strategy from all other strategies available to her. Thus, the player can split the category “any” into the category “focal” and the category “any with two NE”.

This view of the strategy space corresponds to F1 in Figure 2. In the first category, there is only one action - i.e. the focal action, which the player plays with probability 1 if she chooses this category. In the category “any with two NE” there are 14 actions corresponding to the 14 rows (columns). If the player chooses this category she randomizes between the 14 possible actions. Note that the frames are nested. That is, we assume that all categories contained in a lower frame are also contained in higher frames. E.g. the category “any” available under F0 is also available under F1.

A player may further use the fact that there is one row (column) that stands out among those in the category “any with 2 NE”. This is the row (column) of the Associated NE (ANE) (see Figure 1b). Thus, under F2 players can split the category “any with 2 NE” into the two additional categories “associated” and “any other with 2 NE”. Note that the category “associated” can be distinguished only after distinguishing the category “focal”. Again F2 also contains all categories available to a player under the lower frames F1 and F0. The next level of looking at the game is F3. We assume that under F3 a player splits the category “any other with 2 NE” into the 13 rows/columns that it comprises. In principle there can be a number of frames between F2 and the frame that we described as F3. Note that after distinguishing the ANE, a player may realize that there is another NE in the same row (column) as the ANE, and thus she may have a frame that distinguishes the corresponding actions from the remaining rows (columns) with
two NE. Following such steps, eventually all rows (columns) could be distinguished. Note, however, that while the associated action has a clearly distinct feature in that a deviation from the focal to the associated action may lead to a strict payoff gain, there is no such payoff gain from further deviations, as all NE are equivalent. In other words, the distinction between these other rows (columns) does not appear that prominent. Therefore, we assume that under F3 a player splits the category “any other with 2 NE” directly into the 13 rows/columns it comprises.

Following the same logic in the control treatment, i.e. the treatment without the focal point, there are only two possible categorizations. Under F0 a player puts all available strategies in the category “any” and chooses uniform randomly among them. Under F1 a player distinguishes all 15 rows/columns, placing each one of them in a separate category. Note that what we refer to as F1 in the focal point treatments is thus different from F1 in the control treatment.

Figure 2: Frame levels in the BDV game

We now analyze the consequences in terms of players’ expected payoffs from the different possible frame combinations in this game. In a given strategic situation players may use the same frame or they may use different frames. The cases when both players use the same frame are represented in Figure 3. The cases when the players use different frames are represented in Figure 4. If both players use F0, they put all 15 strategies in the same category. Each player randomizes among the

---

This static analysis of the game is in the spirit of VFT as in Bosch-Domènech and Vriend (2013).
15 strategies and her expected payoff is equal to $\frac{1}{15} \times \frac{1}{15} \times (100 \times 29 + 87) = 13.28$.\footnote{We are considering expected payoffs in the 87 treatment. The expected payoffs in the other treatments can be calculated in the same way.}

This is shown in the upper leftmost cell of Figure 3. If both players use F1, each of them has three categories of strategies: “any”, “focal”, “any with 2 NE”. There are two symmetric NE in expected payoffs if both players use F1: (“focal”, “focal”) and (“any with 2 NE”, “any with 2 NE”). The NE in which both use the focal category is Pareto superior to the NE in which both choose the category “any with 2 NE”. Next, we consider the case of both players using F2. Each player has five categories of strategies and there are three NE in expected payoffs, two asymmetric and one symmetric: (“focal row”, “associated column”), (“focal column”, “associated row”), (“any other row with 2 NE”, “any other column with 2 NE”). The two NE in which one player plays the focal category and the other the associated category correspond to the ANE in the original payoff matrix. They are the Pareto superior NE in the case when both players use F2. Finally, Figure 3 shows the case when both players use F3. If both players use F3 there are 29 NE just as in the original normal form payoff matrix: the 2 ANE plus the other 27 NE. All of these NE are equally efficient and entail a payoff of 100.

Figure 4 illustrates the matrices of expected payoffs for the cases when the two players use different frames. The NE are again highlighted in grey. Applying the idea of categorization of strategy spaces to the underlying coordination game presented in Figure 1b, one can think of a transformed game in which players first choose a frame level (categorization), then choose a category within that frame level with any strategy choice within a category being uniform random. The complete normal form of this game corresponds to the case when both players use F3.

The point of Figure 3 and Figure 4 is to illustrate that players’ perceptions of the game may affect their choices as different frame combinations give rise to different NE in expected payoffs. There is a multiplicity of NE if one allows for all possible combinations of frames. Due to the nestedness of the frames, the case when both players use F3 contains all possible combinations of categorizations of the strategy space that players may use. Each of the other frame combinations is Figure 3 and in Figure 4 is an excerpt from this complete payoff matrix.

4.2 Game of Blume and Gneezy (2010)

We now introduce the game used in the experiment by Blume and Gneezy (2010). In this game two players are randomly and anonymously matched. Each player has to independently choose one of the five sectors of the disc illustrated in Figure 5a. If both players choose the same sector, each of them gets a payoff which we
normalize to 1 for our convenience, otherwise they get nothing. The game is played in the following way. The experimenter brings the disc to one of the players first. The player chooses a sector of the disc and the experimenter places a sticker on the corresponding sector on the inside of the disc, i.e. in a way such that the label is invisible on the outside. The experimenter then brings the disc to the second player and the second player makes her choice without knowing the choice of the first player. At the beginning of the experiment the players are informed that the second player may be viewing a rotated disc and that there is a fifty percent chance that the experimenter flips the disc before bringing it to the second player. In a variation of this game, subjects play the game against themselves, following

Figure 3: Symmetric frame combinations in the BDV game
Figure 4: Asymmetric frame combinations in the BDV game

The standard normal form representation of the game is shown in Figure 5b. Under this representation there are five equivalent NE, corresponding to both players choosing the same sector. Thus, players face an equilibrium selection problem.

We now discuss different possible representations of the strategy space in the disc game. They are illustrated in Figure 6. The coarsest possible view of the strategy space is the categorization under which a player puts all strategies in one category, the category “any”. This is F0. The sectors of the disc differ in colour. If a player pays attention to colour, she will recognize that there are black and there are white sectors, and she can form two additional categories of strategies - “any black” and “any white”. We call the resulting categorization of the strategy space F1. The structure of the game is such that after distinguishing colour, a player could distinguish an additional attribute. Namely, she could notice that even if the disc is flipped or rotated, the relative position of one of the white sectors is the same. This is F2.
sectors between two black sectors will be preserved. This is not the case for the other two white sectors or the other two black sectors, which are indistinguishable from each other if the disc is flipped. This property, which Blume and Gneezy
Figure 7: Symmetric frame combinations in the BG game

(2010) refer to as flip symmetry, makes the white sector between the two black
a) $F_0 \times F_1$

<table>
<thead>
<tr>
<th>any</th>
<th>any black</th>
<th>any white</th>
<th>distinct white</th>
<th>any other white</th>
</tr>
</thead>
<tbody>
<tr>
<td>any</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td></td>
</tr>
</tbody>
</table>

b) $F_0 \times F_2$

<table>
<thead>
<tr>
<th>any</th>
<th>any black</th>
<th>any white</th>
<th>distinct white</th>
<th>any other white</th>
</tr>
</thead>
<tbody>
<tr>
<td>any</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>any black</td>
<td>0.2</td>
<td>0.5</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>any white</td>
<td>0.2</td>
<td>0.0</td>
<td>0.3</td>
<td>0.3</td>
</tr>
</tbody>
</table>


c) $F_0 \times F_3$

<table>
<thead>
<tr>
<th>any</th>
<th>any black</th>
<th>any white</th>
<th>distinct white</th>
<th>any other white</th>
</tr>
</thead>
<tbody>
<tr>
<td>any</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>any black</td>
<td>0.2</td>
<td>0.5</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>any white</td>
<td>0.2</td>
<td>0.0</td>
<td>0.3</td>
<td>0.3</td>
</tr>
</tbody>
</table>

d) $F_1 \times F_2$

<table>
<thead>
<tr>
<th>any</th>
<th>any black</th>
<th>any white</th>
<th>distinct white</th>
<th>any other white</th>
</tr>
</thead>
<tbody>
<tr>
<td>any</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>any black</td>
<td>0.2</td>
<td>0.5</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>any white</td>
<td>0.2</td>
<td>0.0</td>
<td>0.3</td>
<td>0.3</td>
</tr>
</tbody>
</table>


e) $F_1 \times F_3$

<table>
<thead>
<tr>
<th>any</th>
<th>any black</th>
<th>any white</th>
<th>distinct white</th>
<th>any other white</th>
</tr>
</thead>
<tbody>
<tr>
<td>any</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>any black</td>
<td>0.2</td>
<td>0.5</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>any white</td>
<td>0.2</td>
<td>0.0</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>distinct white</td>
<td>0.2</td>
<td>0.0</td>
<td>0.3</td>
<td>0.0</td>
</tr>
<tr>
<td>any other white</td>
<td>0.2</td>
<td>0.0</td>
<td>0.3</td>
<td>0.0</td>
</tr>
</tbody>
</table>

e) $F_2 \times F_3$

Figure 8: Asymmetric frame combinations in the BG game

Sectors unique. A corresponding internal representation of the strategy space is $F_2$. 

17
in Figure 6. Distinguishing the attribute flip-symmetry additional to the attribute colour adds two categories of strategies in F2 - the “distinct white” category and the “any other white” category. As before, F2 contains also the categories from the preceding frames. Furthermore, a player who has the disc in front of her in a particular position could in principle label the two black and the two other white sectors for herself, e.g. the first black sector clockwise may be “b1”, the second black sector clockwise “b2”, the first white sector clockwise “w1”, and the second white sector clockwise “w2”. This categorization corresponds to F3. The labels “w1”, “w2”, “b1”, and “b2” are private labels that the agent can assign to the sectors. They are not available in the common language of the two players. Note that they are even not available if the player plays a game against herself. That is, if the experimenter tells her to choose a sector and then takes the disc away from her and returns it without the player knowing whether the disc has been flipped, then the player has no way of distinguishing “w1” from “w2”, and “b1” from “b2”.

We now present the expected payoffs under the different possible combinations of frames and the NE that arise under each combination. First, we consider only symmetric combinations of frames, i.e. combinations of frames such that both players use the same frame. They are shown in Figure 7.

We begin by considering the expected payoffs if both players use F0. Using F0 means that players place all five sectors in the same category and choose uniform randomly among them. The expected payoff of a player is equal to the probability that the two players both select a specific sector times the payoff they would get from selecting the same sector. Thus, the expected payoff from both using the category “any” under F0 is equal to \( \frac{1}{5} \times \frac{1}{5} \times 5 = 0.2 \) (payoff in case of coordinating is normalized to 1). This is illustrated in the first part of Figure 7. Next, we consider the expected payoffs if both players use F1. Under F1 each player has three available categories of strategies: “any”, “any black”, “any white”. There are three symmetric NE in expected payoffs if both players use this representation of the game. These involve the category profiles (“any”, “any”), (“any black”, “any black”), and (“any white”, “any white”), respectively. That is, all cases in which players use the same category of strategies are NE. Note that if both players use F1 the Pareto-superior NE in expected payoffs is (“any black”, “any black”), as in that case both players get an expected payoff of 0.5 compared to an expected payoff of 0.3 in the NE (“any white”, “any white”) and an expected payoff 0.2 in the NE (“any”, “any”). Now consider the matrix of expected payoffs if both players use F2. Each player has five categories of strategies: (“any”, “any black”, “any white”, “distinct white”, “other white”). Under this frame combination there are five NE, again corresponding to both players choosing the same category of strategies. The Pareto superior NE in expected payoffs is the category profile under which both players choose the category “distinct white”. 

18
The last matrix of expected payoffs in Figure 3 is the matrix of expected payoffs resulting from both players using F3. Under F3 each player has nine different categories of strategies. The additional categories of strategies under F3 compared to F2 are the categories “b1”, “b2”, “w1”, “w2”. Note, however, that the expected payoffs from the category profile (“w1, w1”) are the same as the expected payoffs from the category profile (“any other white, any other white”). The same is true about the category profile (“w2, w2”). The category profiles (“b1, b1”), (“b2, b2”), and (“any black, any black”) are also payoff equivalent.

This is because if the disc is flipped players have no actual way to distinguish between “w1” and “w2”, and between “b1” and “b2”, respectively. The two other white sectors and the two black sectors are flip symmetric. Thus, when looking at the disc the player can give the sectors private labels such as “b1”, “b2”, “w1”, “w2”. However, in the language of Crawford and Haller (1990), these are not attainable strategies for the players, as they would not be able to distinguish “w1” from “w2” if they do not know whether the disc is flipped or not. In Sugden (1995)’s terminology the players do not possess common knowledge of the private labels that they assign to the strategies. Therefore the expected payoffs from the NE that involve these category profiles are the same as the expected payoffs from the NE that involve the category profiles (“any black”, “any black”) and (“any other white”, “any other white”). Thus, although in the case when both players use F3, there are additional NE compared to the NE under previous frame combinations, none of them is Pareto superior to the NE (“distinct white”, “distinct white”).

So far we discussed the expected payoffs in the case when the two players use symmetric frames. It is also possible that one player views the game differently from the other. Figure 8 therefore presents the matrices of expected payoffs from different category profiles under asymmetric frame combinations. We note again the prevalence of a multiplicity of NE under the various frame combinations in Figures 7 and 8. One can, again think of the $F3 \times F3$ (Figure 7) frame combination as the complete normal form representation of the transformed game in which players choose first a frame level and then a category. The relative attraction of various strategies depends on the players’ perception of the game with the existence of NE and their Pareto-ranking depending on this perception. The NE in which both players choose the category “distinct white” is Pareto superior to all NE under all possible combinations of F0, F1, F2, and F3 (both symmetric and asymmetric).
5 Model Analysis

The goal of our model is to gain some further understanding into the kind of behavior we may expect in the two games presented in the previous section, both in terms of the views of the strategy space players choose and in terms of outcomes realized. In this section we first confront our model with the data from the BDV experiment for which we have a rich experimental dataset. We then further analyze the dynamic properties of the model for the BDV game, e.g. its aggregate long run payoff maximizing behavior. Finally, we also apply the model to the BG game, where we focus on long term payoff maximizing behavior.

5.1 Model Analysis for Bosch-Domènech and Vriend (2013)

In this section we analyze our model in relation to the BDV experiment. The focus is on three questions. Our first goal is to use the model to generate potential explanations of the behavior observed in the experiment. In the experiment we observe the extent of coordination and the outcomes that people coordinate on. We use these data to find the parameters of the model that generate behavior that most closely matches the behavior observed in the experiment. What we do not observe in the experiment are the private views of the strategy space that players use to make their choices. We therefore analyze the dynamics of use of categorizations and categories in our model under the parameters that best fit the experimental data to gain some insights into the question which views of the strategy space are compatible with behavior observed in the experiment. Second, in the experiment participants play the game for 50 rounds. We employ the model to make some predictions about players’ long term behavior. That is, we consider which outcomes would be realized in the long term, as well as which categorizations and categories players would use. Finally, we take the perspective of a social planner and we further analyze the dynamic properties of our model, by searching for the parameter values that would maximize players’ aggregate long term payoffs. Again we analyze the long term outcomes, as well as the views of the world that emerge, and we compare these findings with the behavior of the model fitted to the experimental data.

The BDV Experiment

We first briefly summarize some key features of the BDV experiment. The main result of the experiment is that the existence of a non-equilibrium focal point can enhance coordination in the game with 29 payoff-equivalent NE described in section 3.1. To show the effect of a non-equilibrium focal point on coordination,
BDV compare behavior in a control treatment in which there is no focal point with behavior in several alternative treatments with a focal point. The treatment variable in the focal point treatments is the magnitude of the focal payoff - 46, 87, and 99, respectively. BDV also analyze how the magnitude of the focal payoff affects coordination. They hypothesize that the higher the focal payoff, the greater the coordination achieved. And that enhanced coordination in treatments with higher focal payoff is achieved either through coordination on the focal point or on one of the ANE.

In the experiment there were 6-8 sessions for each treatment, with 18 participants in each session. They were randomly divided into an equal number of row and column players at the beginning of the session. They kept their roles throughout the session. The game was played for 50 rounds with random rematching each period. The only feedback that the participants received after each round was the payoff from the interaction. They did not receive further feedback on the other player’s action.

The experimental data on average expected payoff per player per period per outcome for each treatment is presented in Figure 9.10 Conditional on coordinating, there are four possible types of outcomes in the BDV game: coordination on the focal point, coordination on the 1st ANE, coordination on the 2nd ANE, coordination on any other NE.11 We observe the following stylized facts in the experiment. First, as Figure 9 shows, the average expected payoff increases over time in all three treatments with a focal point, indicating that people learn to coordinate. Its change in the control treatment is almost negligible. Thus, the existence of a focal point enhances coordination in the experiment. Second, the higher the focal payoff, the higher the average expected payoff players achieve towards the end of the 50 periods. That is, the average expected payoff towards the end of the fifty periods is higher in the 99 than in the 87 than in the 46 treatment, with the difference between the 87 and the 46 treatment being more pronounced. Third, the increase in average expected payoff in the focal point treatments is driven either by increased coordination on the focal point or by increased coordination on an ANE or by both. More precisely, in the 46 treatment it is mainly the increase in coordination on the ANE that drives the increase in average expected payoff over time, while coordination on the focal point stays more or less constant. In the 87 and 99 treatments both coordination on the focal point and coordination on an ANE increase over time. We observe that the lower

---

10The authors report expected rather than actual payoff in order to clear the data of pure random matching effects. The expected payoff thus is based on all pairs that could have been formed rather than those that were actually formed.

11Note that in a given experimental session one ANE is usually by chance realized more often than another. Which one this is varies from session to session. The label “1st ANE” refers to the ANE that is realized more often in a given session.
the focal payoff, the higher the frequency of coordination on the ANE relative to coordination on the focal point.

![Figure 9: Average expected payoffs in BDV experiment](image)

Fitting the Model to the Experimental Data

As our first goal is to use the model to enhance our understanding of the behavior in the experiment we start by estimating the model parameters that give the best fit between the behavior observed in the experiment and the behavior generated by the model. The indicator we focus on is the average expected payoff per player per period depending on the outcome realized, i.e. the indicator that we just presented experimental data on. This indicator captures two important aspects of the dynamics of coordination. On the one hand, it takes into account the frequency of coordination. On the other hand, it shows the relative frequency of different outcomes conditional on coordinating. Thus, it accounts for the extent of coordination and it measures, given that people coordinate, which outcomes they coordinate on.

We conduct a grid search of the parameter space looking for the model parameters that give the closest fit between average payoff per player per period per outcome observed in the experiment and the corresponding indicator generated by our model. Closest fit is measured as the minimal mean squared error between experimental data and model behavior over the 50 periods. We base the grid search on the data from all four treatments giving equal weight to each treatment. Thus, for each period for each outcome we compute the mean squared error between the experimental data and the model. We then sum this mean squared error over all outcomes, all periods, and all treatments. And we look for the model parameters that minimize this sum.
As we explained above, there are three free parameters in our model, and to keep the model simple we decided to use only one $s_0$, one $\beta$, and one $\alpha$ for all choices, players, and periods. In addition we use the same parameter values for all treatments, including the control treatment. We conducted a $10 \times 10 \times 10$ grid search for the values of the parameters that best fit the experimental data.\footnote{For $s_0$ we considered all values in the range $[0.1, 1]$ in increments of 0.1. For beta, all values in the range $[0, 45]$ in increments of 5, and for $\alpha$ all values in the range $[0.1, 1]$ in increments of 0.1.}

The mean squared error between the experimental data and the model results is minimized by the following parameter set: $s_0 = 0.1$, $\beta = 20$, and $\alpha = 0.7$. Thus, the best fit to the experimental data is given by a low initial strength, and a high learning rate (0.7 is the weight placed on the last period). Under this parameter set the average error amounts to 8 (out of 100) per period per player.

### Short term Behavior of the Model Fitted to the Experimental Data

We now analyze the behavior of the model under the parameters that best fit the experimental data. We first present the data on average payoff per player per period per outcome for each treatment and compare it to the data from the experiment. We then look at the dynamics of the categorizations and categories that players use in the model. This allows us to form conjectures about possible views of the strategy space that may have generated the behavior observed in the experiment.

The model data on average payoff per player per period per outcome under the optimal fit parameters is presented in Figure 10. The behavior of our model matches the following stylized facts from the experiment. First, as in the experiment we observe that the average payoff increases over time in all three treatments with a focal point, indicating that agents learn to coordinate. The average payoffs generated by our model in the control treatment are relatively constant and stay at a much lower level than in any of the focal treatments. Second, we see that the higher the focal payoff, the higher the average payoff per player per period at the end of the 50 periods, again with the difference between the 87 and the 46 treatments being more pronounced. And third, the increase in average payoff per player per period in the focal point treatments is driven by increased coordination on the focal point and by increased coordination on the ANE, and not by increased coordination on the other NE.

In order to gain some insights into which views of the world are most consistent with behavior observed in the experiment, we present the dynamics of the relative frequency of categorization use. Figure 11 illustrates the relative frequency of frame use over time for all four treatments (averaged over 100 runs). We see that in all focal point treatments players increase their use of F1 and F2 and decrease...
their use of F0 and F3 over time. A difference between the 46 treatment, on the one hand, and the 87 and 99 treatments, on the other hand, is that in the 46 treatment the use of F2 is slightly higher than the use of F1 at the end of the 50 periods, while the opposite is true for the 87 and 99 treatments. In the 100 treatment both F0 and F1 are used about half of the time. As we discussed earlier F1 in the control treatment is the frame under which a player distinguishes all 15 rows/columns and places them in separate categories.

Figure 11: Frame use BDV model short term
The next question we consider is, conditional on a given frame being used, which categories within this frame do players choose? Figure 12 represents category use over time under the different frames for each of the three focal point treatments. The control treatment is omitted as players in the control treatment use all 18 categories more or less equally often, with a relative frequency of under 0.1 each.

We now analyze the category use within each frame in the different focal point treatments. In all treatments the frame F0 contains only the category “any”. Thus, whenever F0 is used, the category “any” is used. Therefore no graph of category use under F0 is presented. Figure 12 shows that in all three focal point treatments whenever players use F1, they use predominantly the category “focal”. The use of the categories “any” and “any 2 NE” under F1 almost disappears in all three treatments. Both under F1 and F2 the category “focal” reaches a higher relative frequency in the 87 and 99 treatments than in the 46 treatment. Conditional on choosing F2, players use predominantly the categories “focal” and “associated” in all three treatments. In all three treatments, under F3 players also use the categories “focal” and “associated” more often than the other categories.

Note that overall the differences in relative frequency of use of different categories in each treatment is more pronounced under F1 than under F2 than under F3. The higher the level of the frame, the more categories it contains. It thus takes players longer to explore the usefulness of different categories.

The analysis suggests several insights on the behavior in the experiment. First, we see that players may be learning to use a different frame or view of the strategy space in different treatments. Note that this happens although we use the same model parameters for the different treatments. We also observe that players’ views of the world remain quite mixed in all focal point treatments, even after 50 periods. The use of the frames F1 and F2 is higher than the use of the frames F0 and of the frame F3 in all focal point treatments. Furthermore, players learn to use frames of intermediate level of coarseness more often than the finest and more often than the coarsest possible frame. Next, we look at the predictions of the model for the long term.
Figure 12: Category use BDV model short term
Long term Behavior of the Model Fitted to the Experimental Data

As we already explained, players in the BDV experiment played the game for 50 rounds. In this section we ask the question, if players were to continue learning in the same way, which categorizations and categories would they learn to use in the long term? Which outcomes would emerge? To gain some insights on these issues we analyze the long term behavior of agents in our model under the parameters that best fit the experimental data.

Again, we present the three indicators - average payoff per outcome per player per period, frames used, and categories used under a given frame. Figure 13 presents the dynamics of average payoff per player per outcome per period for the different treatments (averaged over 100 runs). We report a horizon of 30000 periods.\textsuperscript{13}

![Figure 13: Average payoffs BDV model long term](image)

Figure 13 shows that if players continue to learn in the same way, they attain high coordination success in the long term in the 46, 87, and 99 treatments. In all three focal point treatments they coordinate exclusively on the focal point and on the 1st ANE. A difference between the 46 treatment, on the one hand, and the 87 and 99 treatments, on the other, is in the outcome players predominantly coordinate on. In the 46 treatment, coordination on the focal point gradually subsides, while coordination on the 1st ANE increases. Thus, in the long term in the 46 treatment players coordinate almost exclusively on the 1st ANE. In the 87 and 99 treatments, the average payoff per player per period generated by

\textsuperscript{13}We also run the model for longer horizons but the longer horizons do not bring additional insights and we therefore limit our reporting to 30000 periods.
coordination on the focal point is about twice the average payoff per player per period generated by coordination on the 1st ANE. It is interesting to note that in the long term the average payoff per player per period is higher in the 46 treatment than in the 87 treatment. This is contrary to what we observe in the experiment and in the short run behavior of the model. To explain why this happens we look into the views of the world that players use in the different treatments.

Figure 14: Frame use BDV model long term

Figure 14 provides us with details on the long term use of frames in the different treatments, again under the parameter values that best fit the experimental data. We see that in all the focal point treatments players learn to use predominantly F1 and F2. F3 is also used but to a much lesser extent. F0 disappears in the long term. It is worth noting that in spite of these common tendencies there is nevertheless a substantial difference in the frame use in the 46 versus the 87 and 99 treatments. In the 46 treatment players learn to use predominantly F2 and to a lesser extent F1. In the 87 and 99 treatments, the use of F1 dominates the use of F2. Thus, players in the 46 treatment more often use the frame that distinguishes the “associated” action, while in the 87 and 99 treatments, they more use the lower level frame F1, even in the long term. In the control treatment there are only two frames - F0 and F1 and they are both used equally often.

We now analyze the use of different categories under the categorizations that players use in more detail. Figure 15 shows category use under each frame for the three focal point treatments. The relative frequency of category use under F0 is omitted as conditional on using F0, players can only use the category “any”. We observe that whenever F1 is used in any of the focal point treatments, players
learn to use exclusively the category “focal” under F1. Under F2 and F3 there is a difference in the category use in the 46 compared to the 87 and 99 treatments. In the 46 treatment both under F2 and under F3 players initially use the category “focal” most often, but over time its use decreases and the use of the category “associated” increases. Thus, in the long term both under F2 and under F3 in the 46 treatment players use the category “associated” at a relative frequency of around 0.7 and the category “focal” at a relative frequency of about 0.3. In the 87 and 99 treatments players again use the categories “focal” and “associated” under F2 and F3. However, in these treatments they use the category “focal” more often than the category “associated” (relative frequencies of 0.7 and 0.3, respectively).

So what can we learn from the long term analysis of the model fitted to the experimental data? We find that in the long term players can learn to coordinate successfully in the focal point treatments. The views of the world that they use to coordinate are mixed even in the long term. And they differ in the different treatments. An interesting finding is that players choose to use frames of intermediate levels of coarseness even in our long term horizon. The coarsest frame F0 disappears completely, and they also use the finest frame F3 only less than one tenth of the time in all focal point treatments.

Furthermore, we observe that in the long term the payoffs generated in the 46 treatment are higher than the payoffs generated in the 87 treatment. This is contrary to what we observed in the short term analysis of the model and in the experiment where we saw that the higher the focal payoff, the higher the payoffs at the end of the 50 periods. Our analysis of the dynamics of the frames and categories suggests the following explanation. In the short term players in all focal point treatments use F1 very often as it is the frame containing the least categories. Under F1 players can most easily coordinate on the focal point. There are only two categories under F1, the category “focal” and a category containing all other actions. If players in the pool learn to use the category “focal” under F1, their probability of coordinating is 1. By coordinating on the focal point, players obtain the focal payoff, which is 46, 87 or 99, respectively, in the different focal point treatments. However, if one player deviates to using the “associated” action, they would obtain a payoff of 100. Thus, over time players learn to use F2, as it has a separate category for the “associated” action. In the 46 treatment the focal payoff is much lower than the focal payoff in the 87 and 99 treatments. Therefore, in the 46 treatment the ANE is an even more attractive option for the players relative to the options in the other focal point treatments. Thus, what we observe is that in the 46 treatment players learn the category focal in the short term, but then in the long term they use the focal point as a stepping stone to the ANE. In the 87 and 99 treatments, although the ANE there is also Pareto superior to the focal point outcome, players coordinate to a greater extent on the focal point even in
the long run.

Our dynamic analysis thus complements the static predictions of VFT in this game. It shows that players may not always choose to coordinate on the Pareto-superior NE but that in some situations (e.g. the 87 treatment) players may prefer to coordinate on a non-equilibrium focal point. The model provides an explanation for this. There is a trade off between using a more detailed representation of the game that contains the Pareto-superior NE and using a less detailed representation of the game that helps players to coordinate faster. Players may choose representations of the game under which they cannot coordinate on the Pareto-superior NE in expected payoffs. Whether players coordinate on the Pareto-superior equilibrium in expected payoffs depends on the difference between the Pareto-superior NE payoff and the payoffs under the alternatives.
Figure 15: Category use BDV model long term
Behavior of the Model under the Payoff Maximizing Parameters

Figure 16: Average payoffs BDV model payoff maximizing parameters

So far in our analysis we considered the behavior of agents in our model under the parameters that best fit the experimental data. We now take the perspective of a social planner and we ask the question - which parameter values for $s_0$, $\alpha$, and $\beta$ should players use if they want to maximize long term payoffs. Which outcomes would emerge? And which categorizations and categories will be used? We therefore run a grid search for the parameters of our model that would maximize long term total payoff across all treatments and we analyze the behavior of the agents in our model under these parameters, comparing it to the behavior of the agents in the model under the parameters that best fit the experimental data. The horizon we consider is 30000 periods. We find that the payoff maximizing parameters are $s_0 = 0.3$, $\beta = 20$, $\alpha = 0.2$. Under these parameters the average payoff per player averaged over all periods for each treatment is higher than the respective payoff under the optimal fit parameters. Maximizing payoffs requires a slightly higher initial strength of categories and categorizations and more patience on part of the players. That is, the learning rate under the payoff maximizing parameters places a weight of 0.2 on the most recent observation compared to a weight of 0.7 on the most recent observation under the optimal fit parameters. The $\beta$ is the same in both cases. We now analyze behaviour in the model under these payoff maximizing parameters.

Figure 16 shows the dynamics of average payoff per player per period under the payoff maximizing parameters. We observe that players reach very high

---

$^{14}$ Again we also considered a longer horizon but the results do not change significantly, so we report the shorter horizon.
payoff levels in all four treatments. In the 46, 99, and 100 treatments they approach the maximal possible payoff. In the 87 treatment, they attain a payoff of above 90, however, they fall short of the maximal possible payoff even in the long term. Note that our grid search for the payoff maximizing parameters was based on all four treatments simultaneously rather than performed for each treatment separately. In the 46 treatment, average payoff per player per period is generated almost exclusively by coordination on the 1st ANE, which explains why the payoff approaches the maximal possible payoff in this treatment. In the 87 and 99 treatments, players coordinate both on the focal point and on the 1st ANE. This explains the somewhat lower payoff in the 87 treatment, where the magnitude of the focal point is 87. The fact that the average payoff per player per period quickly reaches the maximal possible level in the 100 treatment may seem somewhat surprising at first, but it can be explained by payoff maximization under reinforcement learning. Under the payoff maximizing parameters, if players coordinate on one of the 30 NE in the 100 treatment, this equilibrium gets reinforced. Thus, in any given run, players coordinate on one of the 30 NE and they continue playing it. Which of these 30 NE they coordinate on is determined by chance.

Figure 17 displays the dynamics of the frames that agents in our model use under the payoff maximizing parameters. In the three focal point treatments they use predominantly F1 and F2. The relative frequency of use of F3 is below 0.05 in all three focal point treatments and F0 disappears in the long run. The main difference between the 46 treatment, on the one hand, and the 87 and 99
treatments, on the other hand is that the use of F2 dominates the use of F1 in the 46 treatment, while the use of F1 dominates the use of F2 in the 87 and 99 treatments. In the 100 treatment players quickly learn to use almost exclusively F1, i.e. the frame that allows them to distinguish the 15 different rows/columns of the payoff matrix.

We now consider category use under the different frames in the four treatments (Figure 18). In all three focal point treatments, if players use F1, they use only the category “focal”. Under both F2 and F3, the use of the category “associated” dominates the use of the category “focal” in the 46 treatment. In the 87 and 99 treatments, it is the use of the category “focal” that dominates the use of the category “associated” both under F2 and F3. This analysis shines some light on the structure of the game - the bigger the difference between focal point and NE payoff, the more the focal point is used as a stepping stone.

How does the long term behavior of the agents under the payoff maximizing parameters compare with behavior of the agents under the parameters that give the best fit to the experimental data? First, the payoff maximizing parameters are a slightly higher initial strength ($s_0 = 0.3$ rather than $s_0 = 0.1$) than the optimal fit parameters. Second, under the goal of maximizing long term payoffs players need to be more patient ($\alpha = 0.2$ compared to $\alpha = 0.7$), i.e. they need to place a lower weight on the most recent experience compared to the optimal fit parameters. Long term coordination success is already high under the optimal fit parameters. Under the payoff maximizing parameters it is even higher. With respect to outcomes realized, the main differences are in the 46 and 100 treatments, in both of which players obtain much higher efficiency of coordination under the payoff maximizing parameters. As we explained earlier, the high efficiency in the 46 treatment is driven by quick convergence to coordination on the 1st ANE. The high efficiency of coordination in the 100 treatment is driven by reinforcement of the NE that players coordinate on by chance. In terms of frame use players learn which frames to use quicker under the payoff maximizing than under the optimal fit parameters.
Figure 18: Category use BDV model payoff maximizing parameters
5.2 Model Analysis for Blume and Gneezy (2010)

In this section we apply our model to gain further understanding of the game in the BG experiment. We first briefly summarize the relevant data from the BG experiment. We focus on two treatments - the Partner-Separated treatment and the Self-Separated treatment (which we will refer to as Partner and Self treatment from now on). In the Partner treatment players are randomly and anonymously matched to play the game with another participant. In the Self treatment a participant played the game against herself. The relative frequencies of the different choices in the two treatments are shown in Table 1. Note the following stylized facts. First, people do not exclusively coordinate on the distinct sector in either treatment. Second, the overall coordination success as measured by expected payoffs is higher in the Self than in the Partner treatment. Third, in particular the amount of coordination success through the distinct white sector is higher in the Self than in the Partner treatment. Fourth, in the Self treatment most of the coordination success comes through the distinct white sector, whereas in the Partner treatment it comes from coordination on a black sector.

Table 1: Choice Frequencies in the BG experiment

<table>
<thead>
<tr>
<th></th>
<th>distinct</th>
<th>any black</th>
<th>any other white</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self</td>
<td>0.60</td>
<td>0.12</td>
<td>0.28</td>
</tr>
<tr>
<td>Partner</td>
<td>0.37</td>
<td>0.40</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Notes. These are the relative frequencies of choice in the first round of the Partner-Separated and the Self-Separated Treatments in the BG experiment. The data are taken from Figure 4 on p. 495 in the BG paper.

The static analysis in section 3.4.2 showed the NE in expected payoffs under the different possible frame combinations in the BG game. Among the multiplicity of equilibria in expected payoffs in this game the Pareto superior NE is the one of both players choosing the “distinct white” category. The experimental data we just presented shows that participants do not always choose the “distinct white” sector. Blume and Gneezy (2010) offer two possible explanations for this. On the one hand, players may be simply unable to recognize the uniqueness of the distinct sector. On the other hand, players may be able to recognize it but may believe that the other would not be able to recognize it or that the other believes that they are not able to recognize it and so on. The empirical findings of the experiment present

\[^{15}\]Subjects in the experiment played the game for two rounds. We focus on the data in the first round, as players did not receive any feedback on payoff between the first and the second round.
some evidence in favour of both explanations. The fact that only sixty percent of the players choose the distinct white sector in the self treatment indicates that many people may not be able to recognize the uniqueness of the sector. The higher relative frequency of choice of the distinct white sector in the Self versus the Partner treatment indicates that beliefs about others may also play a role.

Behavior of the Model under the Payoff Maximizing Parameters

We now analyze the BG game with the proposed model. We take the perspective of a social planner and ask the question which parameters of the model maximize players’ long term payoffs. What would be the outcomes realized under long-term payoff maximizing behavior and which views of the strategy space lead to such outcomes? We run a grid search for the set of parameters that maximize total payoff over 10000 periods. Again to keep our model simple we conduct the grid search simultaneously for the Partner and the Self treatments, giving an equal weight to each treatment. We find that the optimal parameters are initial strength of $s_0 = 0.1$, $\beta = 20$, and $\alpha = 0.3$. That is, to maximize long term payoff in the BG game, players need to assign a relatively low initial strength to the categories and categorizations and to update relatively slowly, placing a weight of 0.3 on the most recent observation. Under these parameters the average payoff per player per treatment per period averaged over 100 runs of 10000 periods is 0.971 (out of a maximum possible of 1). For the case when a player plays against others it is 0.935, and for the case when she plays against herself it is 0.996. Note that the parameters that maximize long-term payoff in the BG game are similar to those that maximize long-term payoff in the BDV game (BDV: $s_0 = 0.3$, $\beta = 20$, $\alpha = 0.2$).

Playing against others

We first present our analysis of the model for the case when players play against others. There are 30 agents in the model. Each period all agents are randomly matched in pairs and play the disc game. Figure 19 presents the average payoff per player per period per outcome for the short term and for the long term horizon.

We observe that in the first 100 periods average payoff per person per period is slightly above 0.2, which is the benchmark in case both players randomize over all sectors. The share of payoffs generated by coordination on the distinct white sector increases with time, although the coordination rate is quite low in the first 100 periods. Towards the end of the 100 periods coordination on the distinct

\[16\] We consider all possible levels of initial strength $s_0$ in the range of $s_0 = [0.1, 1]$ in increments of 0.1, all possible levels of $\beta$ in the range of $\beta = [0, 45]$ in increments of 5, and all possible levels of $\alpha$ in the range of $\alpha = [0, 1]$ in increments of 0.1.

\[17\] For the long term horizon we present the data for the first 5000 out of the 10000 periods in order to show the change in behavior more clearly.
white sector has the largest share in realized payoffs, followed by coordination on other white, and on black. Figure 19b presents the same indicator in the long term. We see that in the long term players learn to coordinate successfully and the average payoff per player per period approaches the maximum of 1. The increase in coordination is driven by the increase in coordination on the distinct sector. Coordination on other sectors becomes almost negligible in the long term.
Which frames do players use in the short term and in the long term? Figure 20 shows the dynamics of the relative frequency of frame use. In the short term players learn to use frames F2 and F3 slightly more often than frames F1 and F0. In the long term the usage of F2 increases even further (relative frequency of two thirds). F3 is used at a relative frequency of one third. Players learn not to use
F0 and F1.

We now consider the usage of categories under the different frames. That is, we look at the question, given that a frame is used which categories under this frame are used? The relative frequency of category use under the different frames is illustrated in Figure 21a for the short term and in Figure 21b for the long term. Under F1 the category “any white” is used most often in the short term, followed by the categories “any black” and “any”. Under F2 the usage of the category “distinct white” increases to above 0.4 over the first 100 periods, while the use of all other categories decreases. Under F3 the category “distinct white” is again used most often (relative frequency of 0.2). Under F2 and F3, which are the only frames used in the long term, the players learn to use exclusively the category “distinct white”.

Thus, under the goal of maximizing long-term payoff, players learn to coordinate exclusively on the distinct sector. Their views of the strategy space remain mixed in the long run, i.e. they use both frames F2 and F3. These are the two frames that contain a separate category for the distinct sector and thus make coordination on it easier. The model suggests that in the long term learning by reinforcement is sufficient for agents to learn to coordinate on the distinct sector. It is interesting to note that the use of F2 is higher than the use of F3. Thus, a categorization of an intermediate level of coarseness is used more often than the most complete possible frame.

Playing against self

Above we have presented the analysis of the model for the case when players are randomly matched to play against each other. We now look at the results for the case when a player plays against self. Figure 22 presents the average payoff per period in the short term and in the long term if a player plays against self. If players play against self they learn to coordinate more quickly. After 100 periods their average payoff per period is above 0.8 (compared with around 0.2 in the case when they play against others). The increase in payoff is driven predominantly by coordination on the “distinct white” sector. A comparison between Figure 22 and Figure 19 also confirms that players much more quickly learn to achieve the highest possible level of efficiency when they play against self than when they play against others. Thus, our model in which players do not form explicit beliefs about others matches the stylized facts from the BG experiment that coordination on the distinct white sector is higher in the Self than in the Partner treatment and that overall coordination success is higher in the Self than in the Partner treatment.

Figure 23 shows the dynamics of frame use in the BG game when players play against self. Players quickly learn to use F2 and F3 and to ignore F0 and F1. F2 and F3 are the more complete frames that enable them to distinguish the distinct
white sector from the other actions. In Figure 24, which shows the dynamics of category use within the different frames when the player plays against self, we see that she quickly learns to use exclusively the distinct white sector both under F2 and under F3. F0 and F1 are not used in the long term.

Figure 22: Average Payoffs in BG Game against Self

Figure 23: Frame Usage in BG Game against Self

The key findings of our application of the reinforcement learning model to the BG game are the following. We observe that the relative frequency of coordination when players are randomly matched with each other tends towards 1 in the long term.
term. The learning occurs predominantly in the first 1000 periods. Over time players learn to coordinate exclusively on the distinct white sector. In order to do so they learn to use the only two frames which allow them to distinguish this sector, F2 and F3, and they learn to discard completely all other frames. The use of F2 is more frequent than the use of F3, which can be explained by the fact
that F2 contains less categories and it takes less experimentation to discover the
distinct white sector under F2 than under F3.

The comparison of our results of the model when players are randomly matched
to play against others with the results when they play against self shows that the
long term outcomes of the two processes of learning are similar. However, when
playing against self players learn much more quickly.

6 Concluding Remarks

This paper considered categorization of the strategy space in games. In a given
game there may be many alternative categorizations of the strategy space. We
presented a simple model of how players choose which categorization and categories
to use in the process of social interaction. The model is modest in terms of cognitive
assumptions. It postulates that players are more likely to choose categorizations
and categories that have performed well in the past. We applied the model to
improve our understanding of the dynamics of behaviour in an experiment by
Bosch-Domènech and Vriend (2013). We find that the model can match the
stylized facts of the experiment and to account for differences between treatments.
We also show that our model can be applied to gain a better understanding of the
game played in the Blume and Gneezy (2010) experiment. Our analysis of the two
games shows that, first, even in the long term players may choose different views
of the strategy space. Second, even if the players have all possible views of the
world available to them, they do not necessarily use the most complete view of the
strategy space. If people need to learn which categorization to use to coordinate,
there may be some tension between fineness and coarseness of categorizations. On
the one hand, learning to coordinate is easier if there are fewer categories, which
favors coarseness. On the other hand, miscoordination within a category is more
likely, the coarser the category is.


References


