Learning Frames*

Vessela Daskalova[†] and Nicolaas J. Vriend[‡]

26 July 2019

Revise & Resubmit at *Journal of Economic Behavior and Organization*

Abstract

Players may categorize the strategies available to them. In many games there are different ways to categorize one's strategies (different frames) and which ones players use has implications for the outcomes realized. This paper presents a model of agents who learn which frames to use through reinforcement. As a case study we fit the model to existing experimental data from coordination games. The analysis shows that the model fits the data well, providing insights into differences between treatments. It identifies trade-offs of using coarse versus fine frames and categories of strategies when it comes to learning.

Keywords: Variable Frame Theory, Coordination games, Categorization, Reinforcement learning, Focal points, Bounded rationality

JEL Classification: C72, C63, C91, D03

---

*We thank colleagues at the CEF conferences in London and San Francisco, at the Amsterdam Symposium on Behavioral and Experimental Economics, at the ASSET meeting in Evora, at the ESA conference in Honolulu, at GREQAM (Aix-en-Provence), at IMEBESS Barcelona, at the Barcelona GSE Summer Forum Computational and Experimental Economics Workshop, at the University of Cambridge, and at the Institute for Advanced Study in Toulouse for comments. All errors are our own. Vessela Daskalova acknowledges IAST funding from the ANR Labex, from ANR under grant ANR-17-EUR-0010 (Investissements dAvenir program), and from Cambridge-INET.

[†]Toulouse School of Economics, Institute for Advanced Study in Toulouse <vessela.daskalova@tse-fr.eu>

[‡]Corresponding author: School of Economics and Finance, Queen Mary University of London, <n.vriend@qmul.ac.uk>

1 Introduction

To choose among different actions, a decision maker first needs a mental representation of the available options. In many cases some attribute(s) of the decision situation may prompt the decision maker to categorize the options available to her. Categorization usually implies that the person treats objects that she has placed in the same category in the same way (Mohlin, 2014; Mengel, 2012a). One important class of situations in which people may categorize is strategic situations. In this paper we consider players who may group their own strategies into categories. Often multiple internal representations of the same situation or multiple frames of it are possible and which one the decision maker chooses has consequences for their behavior and payoffs. Moreover, in many strategic situations which views of the strategy set are best depends on the views of the strategy set chosen by the other players.

An example of how categorizing the strategy set may matter in strategic situations are some coordination games, where it may help players solve difficult equilibrium selection issues. To fix ideas consider the following matching objects example. There are five objects and players have to independently choose one of them. If they both choose the same one, each gets the same positive payoff. If they choose different ones they get a zero payoff each. If there is nothing to distinguish the five objects, the best the agents can do is put them in one category and randomize uniform randomly among them. However, imagine that one of the objects is blue, while the other four are green. If a player pays attention to colour, she may categorize the blue object separately and may form another category in which she puts all the green objects together. If both players do so and make their subsequent choice by selecting the category blue, they would be able to coordinate better than through pure randomization.[1]

The main goal of this paper is to improve our understanding of factors that may determine which frames of the strategy set players choose and which outcomes emerge as a result. We present a model of agents choosing among alternative frames of a given strategy set. Instead of assuming that players have innate cognitive limitations in viewing the strategy set, we assume that ex ante all possible frames of the strategy set are available to each player. Our model is a dynamic model based on reinforcement learning. Players learn which

[1] This is the simplest possible example of a much more general class of situations where many attributes may be present. See, for example, Bacharach (1993); Bacharach and Bernasconi (1997).

frames 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 frames 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 its behavioral properties. Second, we present its application to pure coordination games. A key challenge in these games is the selection of one of a multiplicity of equilibria. Instead of making an a priori assumption of an equilibrium refinement concept, a study of 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 favor one internal representation over another, reasons that may not be obvious in a static analysis.

As a case study we fit the model to empirical data, examining how it may help us gain insights into human behavior in the coordination games from the laboratory experiment by Bosch-Domènech and Vriend (2013). The analysis shows that the model fits the data from the Bosch-Domènech and Vriend (2013) experiment well, accounting also for differences between experimental treatments. We gain the following insights from analyzing the model under the parameters that best fit the experimental data: first, players may coordinate even without using the same view of the strategy set; second, it is not clear that using the most detailed representation of the strategy set is always best; there are both advantages and disadvantages to using a fine representation of the strategy set when we consider learning dynamics. Our model highlights that both the coarseness of the frame and of the category of strategies that players use matter and it helps us understand the trade-offs of using fine and coarse representations when it comes to learning. On the one hand, a coarse frame, i.e. one that splits the strategy set into fewer categories is useful as it makes it easy for a player to find the appropriate category of strategies (fewer options to explore by trial and error). On the other hand, choosing a category that contains fewer strategies within a coarse frame helps players to find the right strategy within the category. The model also helps to understand why a non-equilibrium focal point may be chosen over multiple Pareto-superior Nash Equilibria. The focal point alleviates the equilibrium selection problem players face by prompting them to put it in a separate mental category and thus distinguish it from other available actions.

Taking as a further example the game from an experiment by Blume and Gneezy (2010) we show how the model can be applied to analyze other games for which we do not have rich enough experimental data.

The remainder of this paper is organized as follows. Section 2 discusses some related literature. In Section 3 we present our model. In Section 4 we present a case study of the model. Sections 5 illustrates how the model can be applied to other games. Section 6 concludes.

2 Relation to the Literature

This paper sits at the intersection of several different literatures: 1) on focal points and equilibrium selection in coordination games; 2) on learning in games, in particular reinforcement learning models; and 3) on categorization in games. We now discuss the relation to each of these literatures and to some other relevant papers in the bounded rationality in games literature more broadly.

In terms of the literature on focal points and equilibrium selection in coordination games, our paper is most closely related to Variable Frame Theory (Bacharach, 1993; Bacharach and Bernasconi, 1997; Bacharach and Stahl, 2000), to Casajus (2000, 2001), to Janssen (2001), and to Sugden (1995).[2] A common theme in the above cited papers and in our paper 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 set 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.

The idea that players may have different frames of the strategy set in our model is closely inspired by Variable Frame Theory (Bacharach, 1993; Bacharach and Bernasconi, 1997; Bacharach and Stahl, 2000). Consider our matching object example from the introduction, where players have to select among five objects, one blue, and four green. If one uses the standard normal form to represent this

[2] Earlier work on the topic includes Gauthier (1975) and Schelling (1980).

game there are five payoff equivalent Nash equilibria, corresponding to the two players choosing the same object, and the game poses an equilibrium selection problem. However, if both players view the game under the color frame, 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 underlie 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.[3] 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 the green objects here. 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. While this is perhaps not explicitly stated in these papers, we believe that the Principle of Insufficient Reason and Symmetry Disqualification are closely related to the idea of putting objects in mental categories and we phrase our model in terms of categories of strategies.

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 (equilibrium) outcomes 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 green 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. Casajus (2000) extends some of the underlying ideas of the above papers to be able to select unique solutions in a broader class of games by considering additionally symmetries of strategic forms based on their frames. Binmore and Samuelson (2006) provide an evolutionary game theory explanation of how focal points may arise and Alberti et al. (2012) present a model of the emergence of salience in recurrent games. 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

[3] An early discussion of the Principle of Insufficient Reason is attributed to Jakob Bernoulli. Later discussions include John Maynard Keynes (Keynes, 2007).

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.

There are two main differences between our work and the Variable Frame Theory related literature. First, as in the other papers, in our model we focus on situations where some attributes are distinguished relatively naturally. Given these attributes, there is a collection of possible representations of the situation, i.e. of different frames of the strategy set. In contrast to the literature, we assume that ex ante all these naturally arising frames 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 in terms of perceiving attributes. We complement this previous literature, by focusing on the question how agents learn to make active use of some frames 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 above 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), which assume that the payoff dominant equilibrium in expected payoffs will be chosen. Instead we follow a dynamic approach based on reinforcement learning with players being more likely to select frames that have been more successful in the past.[4] Rather than postulating that players coordinate on the Pareto superior equilibrium in expected payoffs 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 outcomes emerge in the process of social interaction we are also able to learn more about the advantages and disadvantages of different frames when it comes specifically to learning.[5]

In this respect our paper is related to a second literature, namely on learning models, and in particular models that are based on some kind of reinforcement learning, examples of which include Bush and Mosteller (1951); Roth and Erev

[4]Note that we focus on learning frames of the strategy set 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.

[5]A further minor difference with this 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 frames the other players use.

(1995); Börgers and Sarin (1997); Erev and Roth (1998); Camerer et al. (2004).[6]

What we add to the reinforcement learning literature is the idea of learning not actions, but frames and categories of the strategy set. We also show in the paper how our model compares with a reinforcement learning model without frames and categories. Romero and Rosokha (2019) also consider players who learn to categorize strategies through reinforcement. Their model and applications are complementary to ours in various ways. First, they consider categorization as an adaptive process, rather than as arising from natural attributes of the situation. Second, they focus on a different type of games, the indefinitely repeated Prisoner's Dilemma.

Our reinforcement learning model (as also others in the literature) is more modest in terms of cognitive assumptions than most of the belief-based or mixed models cited. Depending on the task at hand, a different model might be suitable. After presenting the model in Section 3 we discuss reasons why we choose a reinforcement learning model for modeling the issues this paper addresses.

The third literature this paper belongs to is the literature on categorization, and categorization in strategic situations in particular. Recent game theoretic literature has postulated that players may categorize other players (Azrieli, 2009, 2010), they may categorize games (Gibbons et al., 2017; Grimm and Mengel, 2012; Heller and Winter, 2014; LiCalzi and Mühlenbernd, 2019; Mengel, 2012a; Huck et al., 2011; Jehiel, 2005), they may bundle states (Jehiel and Koessler, 2008), nodes other players must move at (Jehiel, 2005; Neilson and Price, 2011), or moves in a game (Jehiel and Samet, 2007) into analogy classes. Daskalova and Vriend (2019) consider players who categorize their own past experiences and are interested in making predictions and coordinating their predictions with one another. Unlike those papers we focus on players categorizing their own strategies and learning which among alternative possible frames of the strategy set to use. A common assumption in the categorization literature deriving from previous literature in psychology is that objects within a category are treated in the same way.

In the same spirit this paper is also related to the literature on using coarse rather than fine categories or on why using less information than one has available can be useful. Papers that show that the use of coarse categories can be optimal

[6]See also Sutton and Barto (2018). This is by no means a comprehensive review of the vast literature on learning, other, less related to our work, learning models include Young (1993) and Arifovic and Ledyard (2011), as well as many others, see review in Fudenberg and Levine (1998).

in some contexts include (Mohlin, 2014; Al-Najjar and Pai, 2014; Mengel, 2012b; Daskalova and Vriend, 2019). Here our focus is on the trade-offs that emerge in using coarse and fine frames and categories of strategies if players have to learn which view of the strategy set is useful.

In terms of papers on bounded rationality in games more generally, Arad and Rubinstein (2012) and Arad and Rubinstein (2018) consider players who think of features of strategies rather than strategies per se. Arad and Rubinstein (2012) is an experimental investigation of the Colonel Blotto game linking it to a multi-dimensional reasoning procedure. Arad and Rubinstein (2018) propose multi-dimensional equilibrium as an alternative to Nash in such situations.

In being about a type of bounded rationality in games, this paper bears some relation to the literature on level-k (Nagel, 1995; Stahl and Wilson, 1994, 1995; Camerer et al., 2004; Costa-Gomes and Crawford, 2006). However, it differs in many aspects from this literature. While in level-k, different levels correspond to differences in strategic sophistication, in our framework different levels correspond to differences in perception of the situation, or more precisely to differences in which attributes players take into consideration when forming categories. Moreover, a general feature of level-k models is that a higher level is a best response to a lower level, whereas this is not a feature of our model. With the exception of Alaoui and Penta (2015), much of the level-k literature treats levels as given. To the extent that Alaoui and Penta (2015) consider endogeneizing levels of strategic reasoning and we consider endogeneizing levels of framing of a strategic situation, there is some commonality.

3 Model

This section describes the general set-up of the model. There is a population of n agents who are randomly (re)matched in pairs each period to play the same underlying game Γ. The underlying game Γ is a simultaneous-move one-shot game and thus has a "standard" normal (strategic) form representation. Each agent has an action set $A = \{a_1, a_2, .., a_m\}$ describing her options in the normal form.[7] An agent may categorize the actions available to them. A category is a subset of the set of actions in the underlying game Γ, i.e. $C \subseteq A$.

Different ways to categorize the action set give rise to different mental

[7] As we consider cases where both players have the same options, we leave out indices for players.

representations or different frames. A frame is a set of categories partitioning the set of actions in the underlying game Γ, that is $F = \{C_1, C_2, ..C_k\}$ such that $\bigcup_i C_i = A$ and $C_i \cap C_j = \emptyset$ for all $C_i, C_j \in F$. The set of frames that are feasible in a given game Γ is denoted by $\Phi = \{F_0, F_1, ..., F_f\}$. We define a relation "refines" denoted by "$\succ$" on the set Φ. A frame $F_{x+1} \succ F_x$ if and only if for each C' in F_{x+1} there exists a C in F_x such that $C' \subseteq C$ with at least one $C' \subset C$.

We focus on cases where categories and frames arise "naturally" from attributes that players distinguish.[8] This means that the agents in our model cannot form categories by arbitrarily combining actions but can only categorise actions on the basis of perceived attributes. Examples of such attributes are the color of the objects in the matching objects game discussed in the introduction, a focal payoff in the underlying game, or labels attached to actions in the underlying game. Thus, attributes that prompt one to categorize one's strategies might be (e.g in the case of a focal payoff) or might not be (e.g. in the case of labels) part of the formal description of the underlying game in the normal form.

We assume that players are able to recognize all possible attributes that are distinguished by the modeler. Thus, for example, if a player needs to choose an object, with the available objects coming in different shapes and colors, we assume that each player does recognize these shapes and colors.[9]

The question we focus on is which views of the strategy set players learn to use, and how this may change as they learn in the process of social interaction. Depending on which attributes an agent takes into account, she may look at the strategy set through different frames and use different categories. We think of the different possible frames and categories as competing for the player's attention, with the attention paid to each frame and category depending on the past success when using them. Some frames and categories may be more successful than others, and this may change over time as all players may be learning which frames to use.

Each period each player makes a decision and learns from its outcome. The timeline in the model is presented in Figure 1. To make a decision she needs to

[8] The related idea of the existence of "natural kinds" has a long history in philosophy, see, for example, John Stuart Mill's *A System of Logic* (1858) or Quine (1969). Aristotle's classification of animals as well as many subsequent taxonomies in the sciences can be seen as examples of classification on the basis of the presence or absence of some attribute.

[9] Modeling how the players have come to recognize attributes as such is beyond the scope of this paper. One possible interpretation is that agents have evolved to recognize certain attributes (Berlin and Kay, 1991). Our model would be complementary to such an evolutionary perspective.

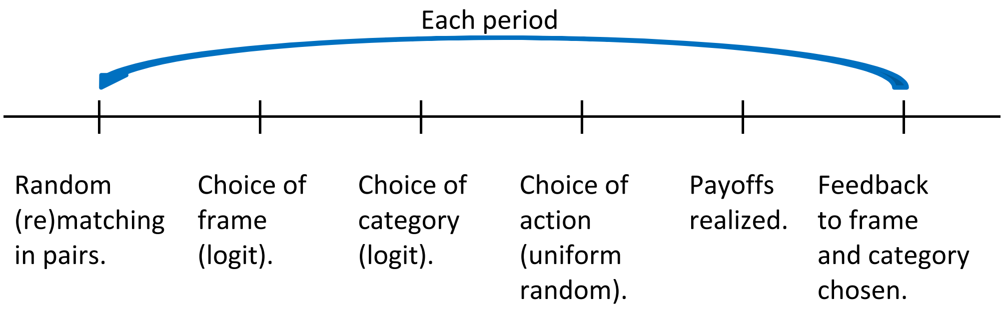


Figure 1: Timeline in the Model

make several choices. First, she needs to choose a frame to use. Then she needs to choose a category from the frame she has selected. And finally, she needs to choose an action from the category chosen. Every frame and category have a measure of how good they are, that is their "strength" as determined by their past performance, and frames and categories are chosen on the basis of these perceived strengths. Actions within a category are chosen uniform randomly.

Thus, we need to specify how the perceived strengths of frames and categories are determined, and how they are taken into account in the choices. We start with the choice of frames. In each period an agent first chooses which frame to use with the logit rule (McFadden, 1974; Cramer, 2003). For a player i the probability of choosing frame j in period t is equal to

$$p_t(F_{i,j}) = \frac{e^{\beta s_t(F_{i,j})}}{\sum_f e^{\beta s_t(F_{i,f})}} \ , \tag{1}$$

where $s_t(F_{i,j})$ is the strength of frame j of Player i in period t and β is a parameter determining the sensitivity of choice to strengths. For $\beta = 0$ choice is uniform random, all frames are chosen with equal probability. If $\beta > 0$ frames that had a better performance in the past are more likely to be chosen.

After choosing a frame to use, a player chooses a category from this frame. This choice is made analogically with the logit rule, where $p_t(C_{i,j,c})$ is the probability that player i in period t chooses category c in frame j.

$$p_t(C_{i,j,c}) = \frac{e^{\beta s_t(C_{i,j,c})}}{\sum_k e^{\beta s_t(C_{i,j,k})}} \ , \tag{2}$$

After choosing a frame and a category, a player chooses an action within the category chosen uniform randomly.[10]

[10] Treating all objects within a category equally is a basic feature of many categorization

We now turn to the determination of the perceived strengths of the frames and categories. The initial strengths of frames and categories are captured by the parameter s_0. Thus, each player starts out with all possible frames and values all of them equally in period 0.[11]

The strength of the frame j that player i uses in period t, $s_t(F_{i,j})$, is updated according to the following rule:

$$s_t(F_{i,j}) = s_{t-1}(F_{i,j})(1-\alpha) + \pi_t \alpha \tag{3}$$

where $s_{t-1}(F_{i,j})$ is the strength of this frame in the previous period, α is the weight the agent places on the latest interaction, and π_t is the payoff from the interaction in period t. The higher the α (with $0 \leq \alpha \leq 1$), the greater the weight the agent places on more recent experiences.

The updating of categories is analogical to the updating of frames. That is, the strength of category c in frame j of player i in period t, $s_t(C_{i,j,c})$, is updated according to:

$$s_t(C_{i,j,c}) = s_{t-1}(C_{i,j,c})(1-\alpha) + \pi_t \alpha \tag{4}$$

Note that some categories may exist in more than one frame of the agent. In that case the strength of a category is updated with the payoff it generated whenever it is used in each frame that it is an element of.

To sum up, there are three free parameters in this model: the initial strength s_0, the learning rate α, and the β in the logit rule. These parameters appear both at the frame and at the category level. In principle the parameter values for s_0, for α, and for β could be different for the frame choice and for the category choice. Furthermore, there could be a specific initial strength for each category and frame. Parameter values could also differ between individual players and they could vary over time. For the sake of parsimony we use for each instance in our model only one parameter value for all initial strengths s_0, only one parameter value for the learning rate α, and only one parameter value for the choice parameter β.[12]

models (Mohlin, 2014; Mengel, 2012b). It is also in line with the Principle of Insufficient Reason.

[11] In principle one could start with any random valuation of the different frames, but it would be difficult to conceptually justify why one would start with one valuation rather than another. We adopt here the Principle of Insufficient Reason, which dictates that absent any reason to treat objects differently one should treat them in the same way.

[12] This implies that we are assuming that all players start out with all possible views of the strategy set and they weigh them in the same way. Also, it implies that they use the same

The agents do not receive any feedback on the frames other players use, on the categories, or on the strategies chosen by them. Neither do they receive feedback on the other player's payoff. They receive feedback only on their own payoff and use this to update the strength of the frames and categories they have chosen. In this sense our model is complementary to some other learning models (such as e.g. fictitious play, or experience weighted attraction learning). There are various reasons to restrict feedback in this model to a player's payoff only. First, it is difficult to envisage how one could justify an assumption of observing another agent's internal representation of the strategy space. Second, giving only own payoff as feedback is consistent with the experiments that we use as a case study. Third, it seems interesting to investigate whether a model as simple as ours can capture the stylized facts observed in an experiment with human subjects.

Related to the limited feedback, the players do not form any explicit beliefs about the frames and categories that other players use. Thus, our approach is complementary to a Bayesian approach where they would have beliefs over others' types, where a type is defined by having a particular frame. Here allowing for explicit beliefs would complicate the model unnecessarily as each agent would need to form beliefs over both the frames and the categories that other players could use, taking also into account that a category may exist under multiple frames.[13]

4 Case Study: The Non-Equilibrium Focal Point Game

4.1 The Game

We first introduce the non-equilibrium focal point game from the experiment by Bosch-Domènech and Vriend (2013) (henceforth BDV), which we use as a case study. The main question this experiment addresses is whether a non-equilibrium focal point can help players coordinate better. At the beginning of

learning mechanism, which does not differ between treatments and does not adapt over time. Thus, ex ante players are identical. However, what is interesting that even so ex post they might learn to use different views of the strategy space.

[13]Note that as the perceived strengths of frames and categories that they use change over time as a result of their interaction with others, players in our model are learning and updating their valuations of how good the different options are.

each experimental session players are randomly assigned to be either a row player or a column player. Each player receives the payoff matrix of the game. Players are randomly and anonymously (re)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 from the control treatment is presented in Figure 2a, where the payoff in each cell represents the monetary reward that is given to each player.[14] There are overall 30 payoff-equivalent Nash equilibria in pure strategies (NE). These NE are scattered in the payoff matrix in such a way 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 or the centre of the payoff matrix).

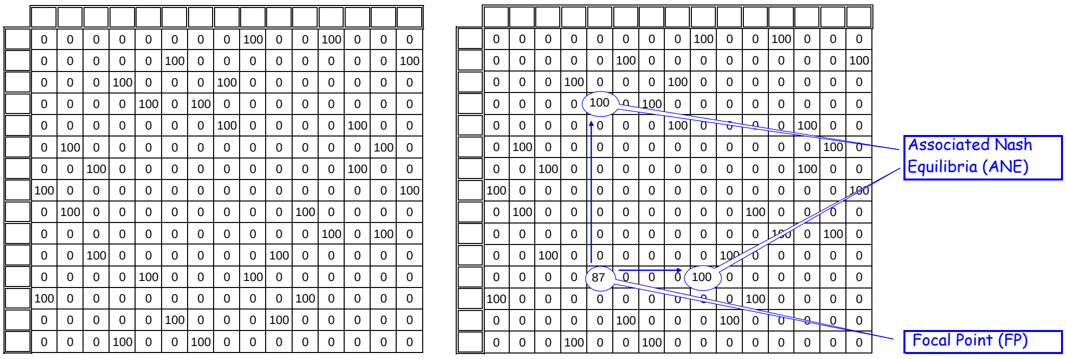
There are three other treatments. In each of them a focal point is introduced by shaving the payoff of one of the NE as in the example in Figure 2b. The three treatments involve a payoff of 46, 87, and 99, respectively. Figure 2b shows the 87 treatment (the payoff matrices for the other two treatments are analogical, but for the magnitude of the focal point). The idea is that 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 an Associated Nash Equilibrium (ANE), i.e. the NE in the same row (column) as the focal payoff. Note, however, that since there are two ANE (one in the same row and one in the same column as the focal point), if both players were to deviate from the focal action in the hope of realizing an ANE, they would miscoordinate. Thus, in the focal treatments instead of 30 NE, there are 29 NE and one non-equilibrium focal point.

We now discuss the set of feasible frames of the strategy space in the non-equilibrium focal point game. In the control treatment (the treatment without a focal point), there are two possible frames of the strategy space. A player can put all possible strategies in one category, the category "any" and then choose uniform randomly among them. We call this frame F0.[15] Or a player could put each strategy in a separate category and have fifteen separate categories, each containing one strategy corresponding to the choice of the respective row (column). We call this frame F1.

Next, we turn to the frames that arise in the treatments with a focal point.

[14]A number of variations of this matrix were used in BDV, consisting basically of rotating or mirroring the payoff matrix, which implies that subjects in different sessions may face slightly differently configured payoff matrices.

[15]In the remaining part of the paper we are slightly abusing the notation we introduced in section 3 and using a number next to F to indicate the level of a frame.



a) without focal point b) with focal point

Figure 2: Game from Bosch-Domènech and Vriend (2013)

They are illustrated in Figure 3. Under F0, in the focal point treatment a player puts all strategies in one category, the category "any". Putting all strategies in one category implies that a player chooses uniform randomly among them. However, a player may choose a more detailed view of the strategy set. 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 set corresponds to F1 in Figure 3. 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 remaining 14 rows (columns). If the player chooses this category she randomizes between the 14 possible actions.

A player may further use the fact that there is one row (column) that sticks out among those in the category "any with 2 NE". This is the associated row (column) (see Figure 2b), including one of the Associated Nash Equilbria (ANE). Thus, under F2 a player partitions 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". 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.[16],[17]

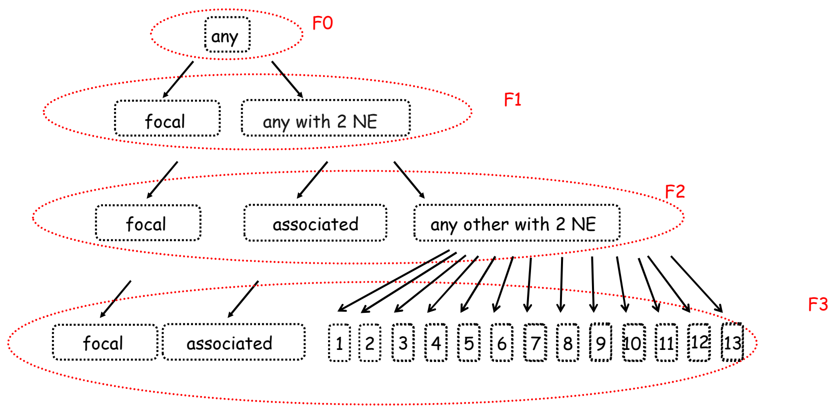


Figure 3: Frame levels in focal point treatments in the BDV game

What are the possible internal representations of the game and their expected payoff consequences when the players apply the frames just discussed? As explained in Sections 2 and 3, we apply the idea that players may categorize the strategy set in line with Variable Frame Theory. This means that, for given frames of the players, the underlying normal form of the game can be transformed into a payoff matrix of expected payoffs from alternative category profiles. Figure 15 in the Appendix shows the consequences in terms of the expected payoffs when both players use the same frame. The cases when the players use different frames are represented analogically in Figure 16. If both players use F0, they put all 15 strategies in the same category "any". Each player randomizes among the 15 strategies and her expected payoff is equal to $\frac{1}{15} \times \frac{1}{15} \times (100 \times 29 + 87) = 13.28$.[18]

[16]In principle there can be a number of frames between F2 and the frame that we described as F3. 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.

[17]Note that the frames are based on attributes of the game that players perceive, and that a higher frame means that players pay attention to more attributes. These levels should not be confused with the levels of reasoning of the level-k literature (Nagel, 1995; Stahl and Wilson, 1994, 1995; Camerer et al., 2004; Costa-Gomes and Crawford, 2006), where players differ in their depth of reasoning in a strategic situation and each higher level of reasoning is a best response to players' using the previous level of reasoning. In our model a higher frame is not necessarily a best response to lower frames.

[18]We are considering here the expected payoffs in the 87 focal point treatment. The expected payoffs in the other treatments can be calculated in the same way.

This is shown in the first table of Figure 15. If both players use F1, each of them has two categories of strategies: "focal" and "any with 2 NE". As we can see, if both players use F1 there are two equilibria: ("focal", "focal") and ("any with 2 NE", "any with 2 NE"). The equilibrium in which both use the focal category is Pareto superior to the one in which both choose the category "any with 2 NE". Variable Frame Theory, using the Principle of Payoff Dominance to select among alternative equilibria, would predict that ("focal", "focal") will be played if both players make a decision under frame F1. Next, we consider the case of both players using F2. Each player has three categories of strategies and there are three equilibria, two asymmetric and one symmetric: ("focal row", "associated column"), ("associated row", "focal column"), ("any other row with 2 NE", "any other column with 2 NE"). The two equilibria in which one player plays the focal category and the other the associated category are the Pareto superior equilibria in the case when both players use F2 and hence the VFT prediction.[19] Finally, Figure 15 shows the case when both players use F3. There are 29 equilibria just as in the original normal form: the 2 ANE plus the other 27 NE. All of these equilibria are equally efficient and entail a payoff of 100.

If we also consider asymmetric frame profiles, we note that from symmetric frame profiles F0xF0 and F1xF1, an individual player has a strict incentive to unilaterally go one frame level higher (in this static perspective). From F1xF1, moving to F2 allows a player to get the ANE payoff of 100 instead of the focal payoff of 87.

4.2 Experimental Data

We now summarize some key findings of the BDV experiment. 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 fifty rounds with random rematching in 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.

Figure 4 shows the experimental data on average expected payoff per player

[19] In contrast, under the extension of VFT proposed by Janssen (2001), the Principle of Symmetry Disqualification would be applied to the two Associated Nash Equilibria as they are payoff equivalent and thus the predictions would be ("any other row with 2 NE", "any other column with 2 NE").

per type of outcome for each period for each treatment.[20] Conditional on coordinating, we can distinguish four types of outcomes in the BDV game: coordination on the focal point, coordination on the 1st Associated Nash Equilibrium (ANE), coordination on the 2nd ANE, and coordination on any other NE.[21] We focus on these payoffs as they capture various important aspects of coordination success, including the frequency and extent of coordination success, as well as the different outcomes players coordinate on.

We observe the following stylized facts in Figure 4. First, the average expected payoff increases significantly over time in all three treatments with a focal point, indicating that people learn to coordinate over time, whereas in the control treatment there is only some almost negligible increase. Second, as a result, in the focal treatments people learn to achieve higher expected payoffs than in the control treatment. Thus, the existence of a focal point helps people to learn to coordinate in the experiment. Third, the higher the focal payoff, the higher the average expected payoff players achieve towards the end of the fifty periods. That is, the average expected payoff towards the end of the fifty periods is higher in the 99 than in the 87 treatment, and higher in the 87 than in the 46 treatment, with the difference between the 87 and the 46 treatments being more pronounced than the difference between the 99 and 87 treatments. Fourth, the increase in average expected payoff in the focal treatments is driven either by increased coordination on the focal point or by increased coordination on the ANE or by both, but not by increased coordination on the other NE. Fifth, the expected payoffs realized through the ANE increase over time in all focal treatments. Sixth, the higher the focal payoff, the higher the frequency of coordination on the focal point relative to coordination on the ANE at the end of the fifty periods. Seventh, the higher the focal payoff, the less frequent the coordination on the other NE.

4.3 Fitting the Model to the Experimental Data

We now fit the model presented in Section 3 to the data from the BDV experiment presented in Figure 4. As we explained above, there are three types of parameters

[20]The 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 in a session rather than just those that were actually formed.

[21]Note 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.

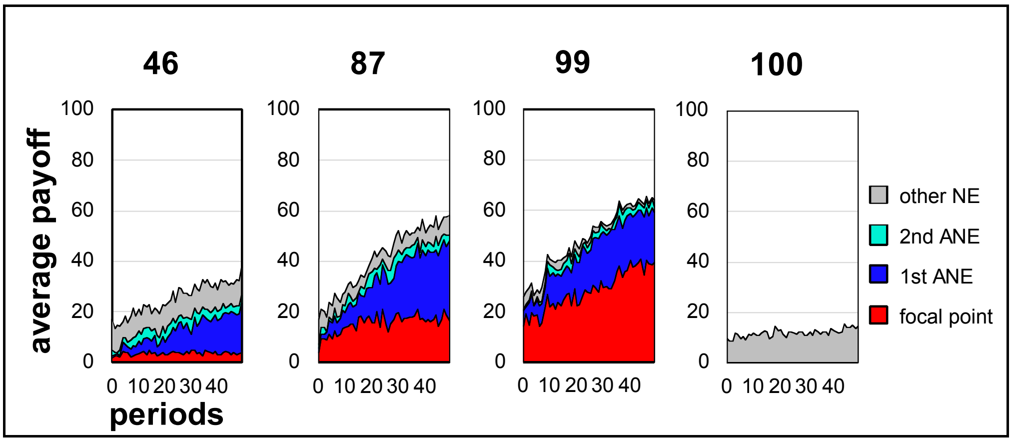


Figure 4: Average expected payoffs per player in BDV experiment

in the model: the initial strength s_0, the temperature parameter in the logit rule β, and the learning rate α. We look for only one value for s_0, one value for β, and one value for α. That is, for each of these parameters we assume the same value across all players, all periods and all decisions (choice of frames and choice of categories), as well as across all treatments (including the control treatment). Thus, our model has only three free parameters. The idea is that if this simple model fits the data well, then any model with more parameters should be expected to fit even better. To look for the model parameter values that best fit the average behavior observed in the experimental data as summarized in Figure 4, we conducted an $11 \times 11 \times 11$ grid search of the parameter space.[22]

For each treatment, we compute the four average expected payoffs per player for each of the fifty periods, distinguishing the expected payoffs per player realized through the focal point, each of the two ANE and any of the other NE, for the experimental data as well as for the data generated by the model.[23] We then compute the mean squared errors between the experimental data and the model data, summing the mean squared errors over the four types of expected payoffs, over all periods and over all treatments, and we look for the parameter values that minimize this sum.

We found that the mean squared error between the experimental data and

[22]Having all payoffs normalized to [0.0, 1.0], for s_0 we considered all values in the range $[0.0, 1.0]$ in increments of 0.1, for β all values in the range $[0, 50]$ in increments of 5, and for α all values in the range $[0.0, 1.0]$ in increments of 0.1.

[23]The experimental data for each treatment are averaged over all (6-8) sessions in that treatment. The model data for each treatment are averaged over 10,000 runs of the model. In each experimental session as well as in each model run, there are nine row and nine column players.

the model data is minimized by the following set of parameter values: $s_0 = 0.1$, $\beta = 5$, and $\alpha = 0.7$.[24]

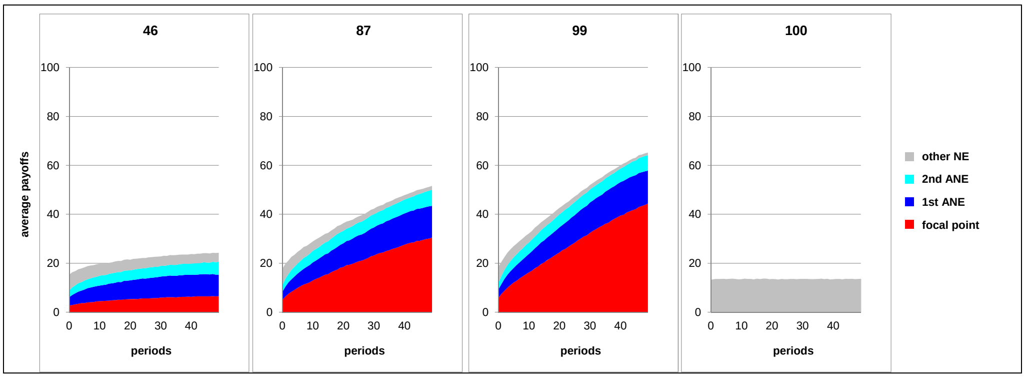


Figure 5: Average payoffs per player learning model

4.4 Analysis of the Model Fitted to the Experimental Data

As Figure 5 shows, the fitted model matches each of the stylized facts of the experimental data listed above. We now analyze the behavior of the model under the parameters that best fit the experimental data. Before considering the dynamics of the choices of frames and categories in more detail, let us provide a short interpretation of the specific parameter values found. The optimal initial strength of 0.1 implies that there will be some exploration of the strategy space as far as the frame and category choices are concerned. But this exploration may be somewhat limited as the strengths of some choices may increase above 0.1. This is also related to the other two parameters. The relatively high $\alpha = 0.7$ suggests that players react relatively fast with respect to their most recent experience as to which frames and categories to use, and that they quickly forget earlier experiences. The $\beta = 5$ means that frames and categories that have performed better in the past have a higher chance of being selected.

In the experiment we observe the actions of the players, the extent of coordination, and the specific outcomes that they coordinate on. What is not observed in the experiment are the views of the strategy set that players use to make their choices. We can use the model to analyze the dynamics of use of frames and categories under the parameters that best fit the experimental data

[24]To check the robustness of this optimization result, we did the grid search 10 times, each time for 10,000 runs, and this is the optimal parameter set for each of these 10 times.

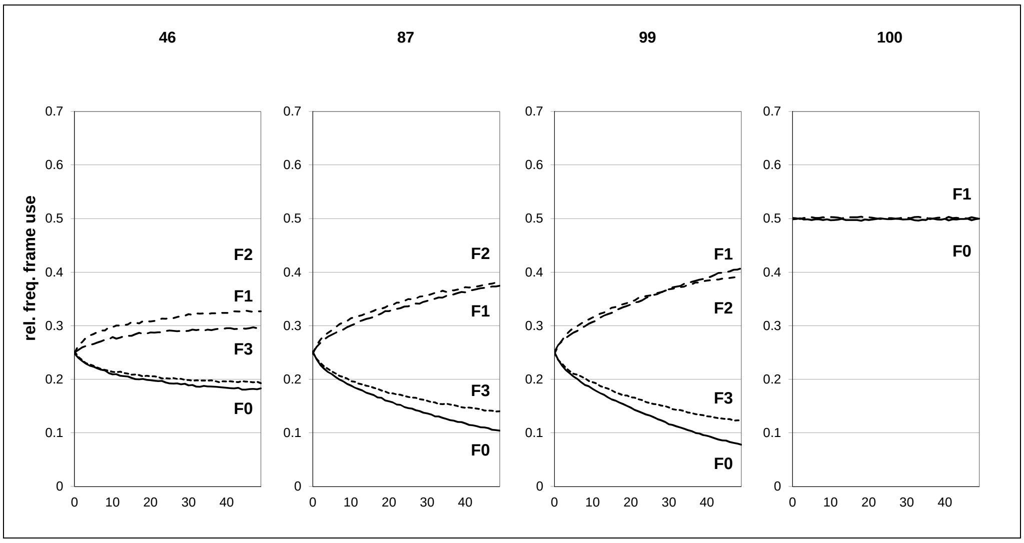


Figure 6: Relative frequency frame use learning model

to gain some insights into the question which views of the strategy set would be compatible with the behavior observed in the experiment.

In order to gain some insights into which views of the game are most consistent with the behavior observed in the experiment, we present the dynamics of the relative frequency of frame use. Figure 6 illustrates the relative frequency of frame use over time for all four treatments (averaged over 10,000 runs). We see that in all focal point treatments players increase their use of F2 and F1 and decrease their use of F0 and F3 over time. In the 46 and 87 focal point treatments F2 is used slightly more often than F1 at the end of the fifty periods, whereas in the 99 treatment F1 is used slightly more often than F2. The two frames in the control treatment are used equally often. Under F0 in the control treatment a player puts all actions in one category and chooses uniform randomly among them. Under F1 in the control treatment each action is put in a separate category and these categories are then reinforced. As there is no difference in expected payoffs from undertaking any action in the control treatment (the payoff matrix contains 30 payoff equivalent Nash Equilibria, two in each row and two in each column), the two views of the strategy space are equivalent when the data is averaged over multiple runs of the model.

We observe that in all three focal point treatments, agents learn the frames of intermediate coarseness levels rather than the most or the least detailed ones. Moreover, the higher the focal point, the steeper the learning curves for these frames over the fifty periods.

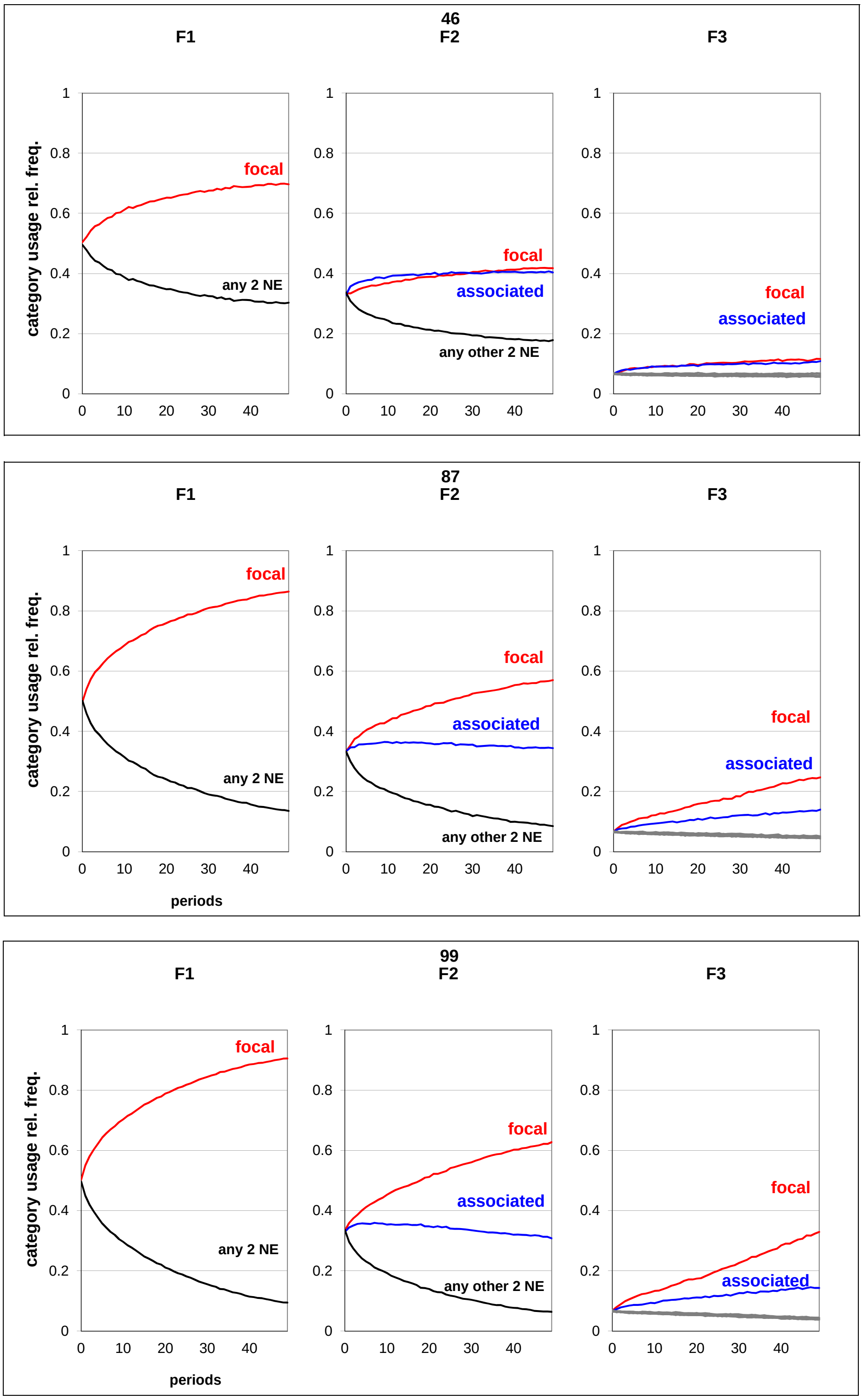


Figure 7: Relative frequency category use learning model

The next question we consider is, conditional on a given frame being used, which categories within this frame do players choose? Figure 7 represents the relative frequency of category use over time under the different frames (F1 to F3) for each of the three focal point treatments. The control treatment is omitted as players in the control treatment under F1 use all 15 categories corresponding to the different rows (columns) more or less equally often, with a relative frequency of about 0.07 each (at the end of the fifty periods).

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 7 shows that in all three focal point treatments whenever players use F1, they use predominantly the category "focal". The use of the category "any 2 NE" under F1 decreases substantially in all three treatments. 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.

In terms of between treatment comparisons, the higher the magnitude of the focal point, the higher the use of the category focal under F1 at the end of the 50 periods. Moreover, both under F2 and F3, the higher the focal point, the higher the relative use of the category "focal" compared to the category "associated" at the end of the 50 periods.

Note that overall the difference in relative frequency of use of those categories whose use increases compared to those categories whose use decreases in each treatment is more pronounced under F1 than under F2 than under F3. The explanation is that the higher the level of the frame, the more categories it contains. It thus takes players longer to explore the usefulness of different categories and they therefore learn more slowly under frames containing more categories.

4.5 Basic Reinforcement Learning without Frames

The analysis above shows that the reinforcement learning model with frames, in which the players learn which frame of the strategy set to use, fits the experimental data well. Before we discuss the possible insights we can derive from using this model in the next section, in this section we consider a closely related reinforcement learning model - without frames. That is, in this model the players use reinforcement learning to respond to feedback about past performance as in our model, but they do not use frames, i.e. they do not categorize the action space. This allows us to assess how important this categorization of the action space is. Would a simpler reinforcement learning model without frames account for the experimental data equally well? Note that a significantly better performance of the model with frames would suggest that the idea of players categorizing the strategy space and learning which frame to use may be relevant as an explanation for the behavior observed in the experiment.

The key difference of this more basic reinforcement learning model with the model proposed in section 3 is that here there are no frames and categories. That is, the agents have only one possible internal representation of the game, corresponding to the standard normal form of the underlying game, i.e. the payoff matrix in Figure 2. Otherwise the model corresponds to our model. Thus, each period two agents are randomly matched to play the game, each agent chooses simply an action, i.e. one of the fifteen possible rows (columns) using the logit rule, payoffs are realized, and the strength of each agent's action is updated with the payoff from the interaction.

In order to compare the two models we carry out a grid search for the parameters of the model without frames that best fit the experimental data. As before, the parameters are the initial strength (here of each action), the β (for choice between the fifteen actions), and the α (to reinforce the strength of each action). And just as before, we minimize the average mean squared error, taking into account the four types of payoffs realized in the experiment (see Figure 4) and those generated by the model. Again the grid search is run 10 times, each time taking the average over 10,000 runs of the model in order to clear out effects due to randomness.

For 6 out of the 10 times, the optimal α found is equal to 0, and for the remaining 4 times, the optimal β found is equal to 0. Note that these findings for the 10 runs are perfectly consistent as these parameter values imply persistent random behavior without any learning. An $\alpha = 0$ means that the strengths of the various choices are not updated on the basis of experience and remain equal to the initial strengths. Since these initial strengths were the same for all choices, the players will continue to choose their actions uniform randomly. A $\beta = 0$ means that the players do not respond to any differences in strength between the various choices and that they make their choices uniform randomly.

This is illustrated in Figure 8, showing the average expected payoffs per player per type of outcome for each period under the optimal parameter values for the

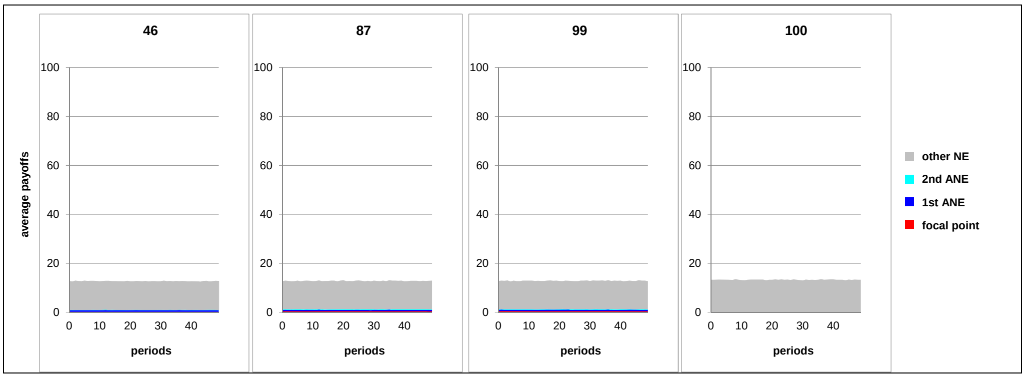


Figure 8: Average payoffs learning model without frames

model without frames. As we can see, in each treatment, including the control treatment, the overall payoffs stay constant over the fifty periods at the level expected for random choices, and the payoffs realized through the focal point or the ANE stay constant at the level expected for random behavior as well, which means that almost all the payoffs are realized through the other NE. In fact, none of the stylized facts observed in the experiment (as illustrated in Figure 4) is captured by this more basic reinforcement learning model in which the players cannot categorize the strategy space, whereas our model (see Figure 5) captured each of them. Thus, the basic reinforcement learning model without frames cannot explain the experimental data any better than uniform random choice of the available actions.[25]

4.6 Discussion

The analysis of the reinforcement learning model with frames suggests several insights into the behavior of the players in the experiment. By modeling frames of the strategy set and the learning of which one to use, we make explicit how

[25]This does *not* mean that agents in a basic reinforcement learning model as such cannot learn to coordinate successfully in a game like the focal point game. This is illustrated in Figure 19 in the Appendix. Instead of fitting the model to the BDV experiment distinguishing the four types of coordination outcomes as in Figure 4, we instead take into account only the overall payoffs realized in each period. As we can see, the overall payoffs per player increase over time in all treatments. The optimal parameter values found are $s_0 = 0.0$, $\alpha = 0.1$ and $\beta = 50.0$. Note, however, that although agents in this reinforcement learning model successfully learn to coordinate, none of the stylized facts observed in the experiment are explained by it. For example, the overall payoff differences between the treatments that we observed in the experiment cannot be explained by this reinforcement learning model, as here the payoff differences between the treatments become negligible, with even the control treatment showing the same increase in total payoffs as the focal treatments, and the focal point and the ANE do not play any significant role.

the players may be viewing the various options presented to them by the focal point game, which attributes they may be taking into account, as well as how all this may be changing over time. In the model these dynamics themselves are governed by the three parameters that we fitted to the experimental data.

What we observe is, first, that agents' views of the strategy set remain quite mixed in all focal point treatments, even after fifty periods. The use of the frames F1 and F2 is higher than the use of the frames F0 and F3 in all focal point treatments. With frame F2 players distinguish the focal as well as the associated action from each other and from all the other possible actions, but besides or instead of this view, often the players learn to use frame F1 in which they distinguish only the focal action from all other actions as the most relevant one. This mixing between different frames may happen at the individual as well as the population level. Thus, what we found is that when two players meet they may actually look in a different way at the strategy set (although the options each of them is facing in the underlying game are the same).

Second, our analysis suggests that while players learn to coordinate more and more successfully, they do not necessarily achieve this by learning finer and finer frames. In fact, our analysis suggests that players may learn to use frames of intermediate levels of coarseness.

Third, we observe that the more detailed the frame of the strategy set the player has (e.g. compare F3 to F2 to F1 to F0), the longer it takes the player to learn which category within this frame to use.[26] This suggests that when it comes to learning, there is some tension between using a more and a less detailed representation of the strategy set.

On the one hand, a lower frame (with fewer categories) has the advantage that the players can learn relatively quickly the value of each category within such a frame, and they can learn quickly on which category to coordinate. On the other hand, a lower frame (with on average more options per category) may have the disadvantage that while coordinating on some category may be facilitated, there will be a good chance of mis-coordination *within* such a category. Thus, from a learning perspective, the best frames are characterized by a mixture of coarseness and fineness: few categories, of which some contain few actions (those on which to base the successful coordination), while the other categories contain many actions. Thus, the analysis of the dynamics of our learning model suggest a rationale for some coarse framing of the strategy set, a rationale that may not

[26]Note that this is true even though experience is carried over from lower to higher frames, i.e. if a category is used by the player it is reinforced under all frames under which it exists.

be apparent in a static analysis.[27]

The dynamic analysis of the reinforcement learning model with frames shows that players may not always learn to coordinate on a (Pareto-superior) NE and that players may learn instead to coordinate on a non-equilibrium focal point (in line with the experimental findings). The model provides an explanation for this, as there is a trade-off between using a more detailed internal representation of the game that contains the Pareto-superior equilibria and using a less detailed representation of the game that helps players to coordinate faster. Players may thus choose representations of the game under which they cannot coordinate on the Pareto-superior equilibrium in expected payoffs (Frame F1). To what extent players coordinate on the Pareto-superior equilibrium in expected payoffs and to what extent they coordinate on the focal point depends on the difference between the Pareto-superior NE payoff and the payoffs under the alternatives. Hence the difference in outcomes observed in the different treatments of the experiment can be understood as resulting from different incentives to learn to use a more detailed frame. The higher the focal payoff (which can be achieved also under F1), the less of an incentive there is to learn to use a more detailed frame (F2).

Besides offering possible explanations as to how players look at the focal point game considered, and how this may change over time, our model may be useful in various other ways. For example, it could be used to predict long-run behavior. That is, given the parameter values fitted to the fifty period experiment, i.e. assuming the players were to continue learning in the same way, one may address the question what these parameter values would imply if the game was played for many more, e.g., one thousand periods, the kind of horizon one cannot normally investigate in an experiment. Which outcomes and what payoff levels would emerge, and which frames and categories would they learn to use in the long term?

One could also use the model to investigate questions of efficiency. For example, what parameter values would maximize the players' payoffs over fifty periods, or in the long run? One could compare these parameter values with those found for the fifty period experiment, and interpret the differences. 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? Which outcomes would emerge? And which frames and categories would be used?

[27]Note that this rationale for using lower frames does not depend on some inability to perceive a more detailed representation of the action space.

Instead of efficiency, one could also use the model to investigate individual optimality. For example, one could fix the behavior of all players using the parameter values found for the model fitted to the experiment, and given those parameter values for all players except one individual player, one could investigate in what sense this individual player could do better. What parameter values would be optimal for this player, how can one interpret any differences with the values found for the experiment, and what does this imply for the outcomes, payoff levels, and frames and categories used?

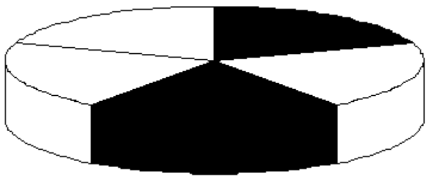
5 An Additional Application. The Disc Game

5.1 The Game

We now introduce the disc game from an experiment by Blume and Gneezy (2010) (henceforth BG), which we also apply our model to.[28] 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 9a. If both players choose the same sector, each of them gets a positive payoff (which we normalize to 1), 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.[29] The standard normal form representation of the game is shown in Figure 9b. Under this representation there are five payoff equivalent NE, corresponding to both players choosing the same sector. Thus, players face an equilibrium selection problem.

[28]Blume and Gneezy's experiment offers some fascinating insights into the respective roles of own cognition and beliefs about cognition in a more or less static context. As we do not have data from more than two periods from the experiment, we will use the game only for the purpose of illustrating how the model can be applied to make prediction in situations for which we don't have sufficient experimental data.

[29]In a variation of this game, subjects play the game against themselves, following exactly the same procedure.



a) Disc

	1	0	0	0	0
	0	1	0	0	0
	0	0	1	0	0
	0	0	0	1	0
	0	0	0	0	1

b) Normal Form

Figure 9: Disc Game (Blume and Gneezy, 2010)

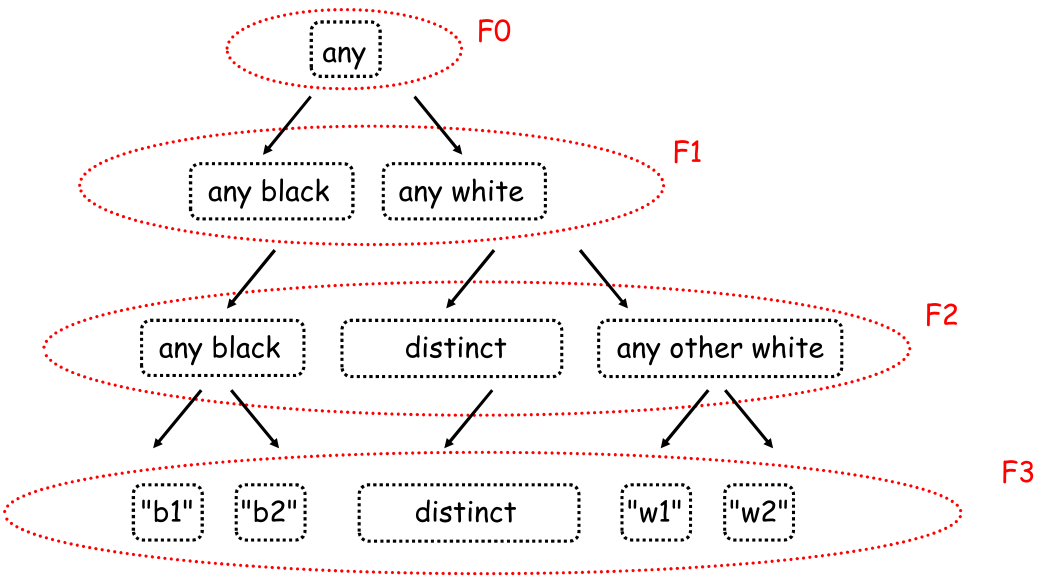


Figure 10: Frame levels in the BG game

There are, however, alternative representations of the strategy set in this game, based on some attributes players could pay attention to. These representations are illustrated in Figure 10. The coarsest possible view of the strategy set is the frame under which a player puts all strategies in one category, the category "any". This is F0. But the sectors of the disc differ in color. If a player pays attention to color, 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 frame F1. The structure of the game is such that after distinguishing color, a player could distinguish an additional attribute. Namely, she could notice that even if the disc is flipped or rotated, the relative position of the one white sector between two black sectors will be preserved. This is not the case for the other two white sectors or the two black sectors, which are indistinguishable from each other if the disc is flipped. This property, which Blume and Gneezy (2010) refer to as flip symmetry, makes the white sector between the two black sectors unique. A corresponding internal representation of the strategy set is F2 in Figure 10. Distinguishing the attribute flip-symmetry additional to the attribute color splits the category "any white" from F1 into two categories of strategies in F2 - the "distinct white" category and the "any other white" category. 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 view of the strategy space corresponds to F3. Here the labels "w1", "w2", "b1", and "b2" are private labels that an agent can assign to the sectors.

We now present the expected payoffs from alternative category profiles under the different possible combinations of frames, focusing on the symmetric combinations of frames, i.e. where both players use the same frame. They are shown in Figure 17 in the Appendix. The asymmetric frame combinations are shown in Figure 18 in the Appendix. 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 times the number of sectors they could coordinate on. Thus, the expected payoff from both using the category "any" under F0 is equal to $\frac{1}{5}\times\frac{1}{5}\times 1\times 5 = 0.2$ (where the payoff in case of coordinating is normalized to 1). This is illustrated in the first part of Figure 17. Next, we consider the expected payoffs under alternative category profiles. If both players use F1, each player has two available categories of strategies: "any black" and "any white". There are two symmetric equilibria in expected payoffs if both players use this representation of the game. These involve the category profiles ("any black", "any black"), and ("any white", "any white"), respectively. That is, all profiles such that the players use the same category of strategies are equilibria. Note that if both players use F1 the Pareto-superior equilibrium 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 ("any white", "any white") equilibrium. As Variable Frame Theory employs the Principle of Payoff Dominance to select among equilibria, this would be the prediction of the theory. Now consider the matrix of expected payoffs under alternative category profiles if both players use F2. Each player has three categories of strategies: ("any black", "distinct", "any other white"). Under this frame combination there are three equilibria, again corresponding to both players choosing the same category of strategies. The Pareto superior equilibrium in expected payoffs is the category profile under which both players choose the category "distinct".

The last matrix of expected payoffs in Figure 7 is the matrix of expected payoffs resulting from both players using F3. Under F3 each player has five different categories of strategies:"b1", "b2", "distinct", "w1", "w2". Note that the expected payoffs from the category profile ("w1, w1") and ("w2, w2") under F3 are the same as the expected payoffs from the category profile ("any other white, any other white"), the latter under F2. The category profiles ("b1, b1"), ("b2, b2"), and ("any black, any black") are also payoff equivalent.[30] 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". But these labels are not available in the common language of the two players. In the language of Crawford and Haller (1990), these are not attainable strategies for the players. In Sugden (1995)'s terminology the players do not possess common knowledge of the private labels that they assign to the strategies.

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 and one could analogically derive the expected payoff matrices under asymmetric frame combinations. Note that, unlike the BDV focal point experiment, in the BG game there is never a strict incentive to choose a higher frame level than the other player, while there is a strict incentive to avoid using one frame level lower than the other player (the exception being where the other player uses F3, as F2 would do equaly well in that case).

[30]Thus, although in the case when both players use F3, there are additional equilibria compared to the equilibrium under previous frame combinations, none of them is Pareto superior to the ("distinct white", "distinct white") equilibrium.

5.2 Experimental Data

We first briefly summarize the relevant data from the BG experiment, focusing on the Partner-Separated treatment. In this treatment players are randomly and anonymously matched to play the game with another participant.[31] The relative frequencies of the different choices are any black 40%, distinct white 37%, and any other white 23%.[32] Figure 11 shows the expected payoffs for the three types of coordination success. Note the stylized fact that people do not exclusively coordinate on the distinct sector.

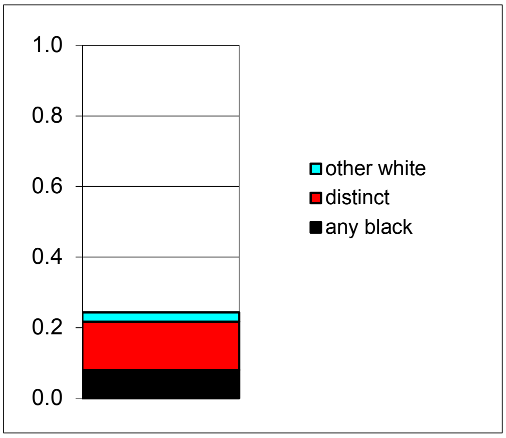


Figure 11: Expected Payoffs in the BG experiment

In the static analysis in Section 5.1 we showed the equilibria in expected payoffs according to Variable Frame Theory under the different possible frame combinations in the disc game. Among the multiplicity of equilibria in expected payoffs in this game the Pareto superior equilibrium 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 some evidence in favour of both explanations. The fact that only sixty

[31] Subjects in the experiment played the game for two rounds, without receiving any feedback on payoffs between the first and the second round. We consider here the data of the first round.

[32] These are the relative frequencies of choice in the first round of the Partner-Separated Treatment in the BG experiment. The data are taken from Table 3 on p. 504 in the BG paper.

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.

5.3 Model Analysis for the Disc Game

We now analyze the disc game with our model. The goal here is not to fit the model to the experimental data[33], rather the goal is to look for the set of parameter values of the model that would maximize the players' fifty-periods payoffs. What are the outcomes realized under payoff maximizing behavior (for a given time horizon), and which views of the strategy space lead to such outcomes?

We run a grid search for the set of parameters that maximize the total payoff over 50 periods.[34] We find that the optimal parameters are initial strength of $s_0 = 0.0$, $\beta = 25$, and $\alpha = 0.3$.[35] That is, to maximize the payoff over a fifty period horizon in the disc 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.

There are 30 agents in the model (as in the BG experiment). Each period all agents are randomly rematched in pairs and play the disc game. Figure 12 presents the average payoff per player per period per outcome.

We observe that even within the relatively short 50 period horizon cumulative average payoff increases, indicating increased coordination over time. The increase in coordination is driven mostly by the increase in coordination on the distinct sector. There is a small increase also in payoffs generated by coordination on the black sectors. The payoffs generated by coordination on the other white sectors are gradually decreasing.

Which frames do players learn to use? Figure 13 shows the dynamics of the relative frequency of frame use. Players learn to use the more detailed frames F2 and F3 more often, whereas the use of the less detailed frames F0 and F1 gradually decreases.

[33] We do not have suitable experimental data to do this.

[34] We consider all possible levels of initial strength s_0 in the range of $s_0 = [0.0, 1.0]$ in increments of 0.1, all possible levels of β in the range of $\beta = [0, 100]$ in increments of 5, and all possible levels of α in the range of $\alpha = [0.0, 1.0]$ in increments of 0.1.

[35] To check robustness we conduct the grid search 10 times, each time over 10000 runs.

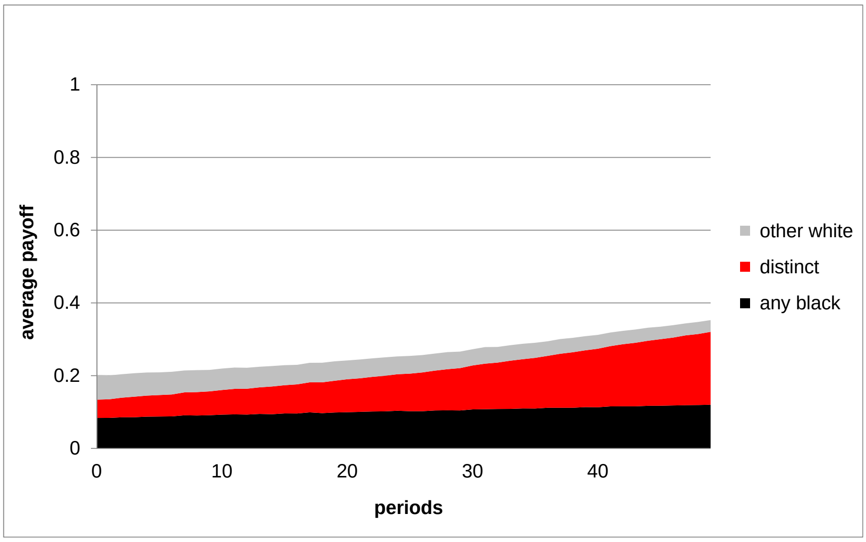


Figure 12: Average payoffs Disc Game

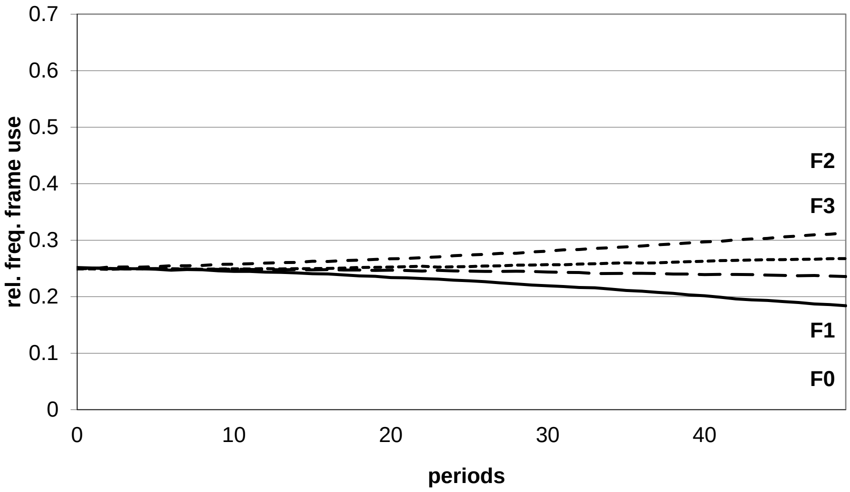


Figure 13: Relative frequency Frame Use Disc Game

We now consider the usage of categories under the different frames. That is, we look at the question, conditional on a frame being used, which categories are used? The relative frequency of category use under the different frames is illustrated in Figure 14. Under F1 players learn to use the category "any black" whereas the use of the category "any white" decreases over time. Both under F2 and F3 players learn to use the category "distinct white" most often. Learning of this category is quicker under F2 than under F3. This can be explained by the fact that under F3 there are more competing categories present than under F2, thus it takes more time to find the "best" category.

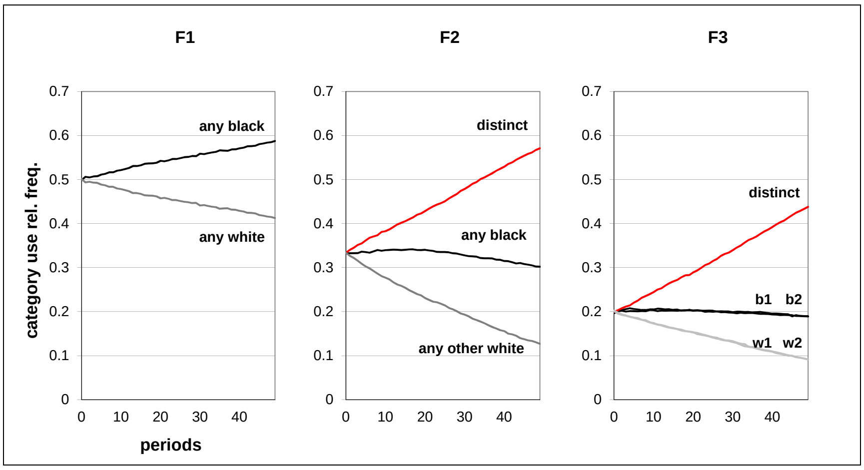


Figure 14: Category use BG model playing against others

6 Concluding Remarks

This paper considered players who categorize the strategy set in some coordination games. In a given game there may be many alternative internal representations of the strategy set and which one players use matters for the outcomes they achieve. We presented a simple model of how players learn which frames 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 frames and categories that have performed well in the past. We applied the model to improve our understanding of the dynamics of behavior in the non-equilibrium focal point game from the 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 possible learning dynamics in games for which we do not have data for a longer time horizon such as the disc 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 set. Second, even if the players have all possible frames available to them, they do not necessarily use the most complete view of the strategy set. If agents need to learn which frames to use to coordinate, there may be some tension between fineness and coarseness of frames. On the one hand, learning to coordinate is easier if there are fewer categories, which favors coarse frames, that is, such containing fewer categories. On the other hand, miscoordination within a category is more likely, the more actions there are within a category.

A Appendix

F0 x F0	*any*
any	13.28

F1 x F1	*focal*	*any with 2 NE*
focal	87	7.14
any with 2 NE	7.14	13.78

F2 x F2	*focal*	*associated*	*any other with 2 NE*
focal	87	100	0
associated	100	0	7.14
any other with 2 NE	0	7.14	15.38

F3 x F3	*focal*	*associated*	*column 1*	*column 2*	*column 3*	*column 4*	*column 6*	*column 7*	*column 8*	*column 10*	*column 11*	*column 12*	*column 13*	*column 14*	*column 15*
focal	87	100	0	0	0	0	0	0	0	0	0	0	0	0	0
associated	100	0	0	0	0	0	0	100	0	0	0	0	0	0	0
row 1	0	100	0	0	0	0	0	0	0	0	0	100	0	0	0
row 2	0	0	0	0	0	0	100	0	0	0	0	0	0	0	100
row 3	0	0	0	0	0	100	0	0	100	0	0	0	0	0	0
row 5	0	0	0	0	0	0	0	0	100	0	0	0	100	0	0
row 6	0	0	0	100	0	0	0	0	0	0	0	0	0	100	0
row 7	0	0	0	0	100	0	0	0	0	0	0	0	100	0	0
row 8	0	0	100	0	0	0	0	0	0	0	0	0	0	0	100
row 9	0	0	0	100	0	0	0	0	0	0	100	0	0	0	0
row 10	0	0	0	0	0	0	0	0	0	0	0	100	0	100	0
row 11	0	0	0	0	100	0	0	0	0	100	0	0	0	0	0
row 13	0	0	100	0	0	0	0	0	0	0	100	0	0	0	0
row 14	0	0	0	0	0	0	100	0	0	100	0	0	0	0	0
row 15	0	0	0	0	0	100	0	100	0	0	0	0	0	0	0

Figure 15: Expected Payoffs under Symmetric Frame Combinations BDV Game

F0 x F1

	focal	any with 2 NE
any	12.47	13.75

F0 x F2

	focal	associated	any other with 2 NE
any	12.47	13.33	13.33

F0 x F3

	focal	associated	column 1	column 2	column 3	column 4	column 6	column 7	column 8	column 10	column 11	column 12	column 13	column 14	column 15
focal	87	100	0	0	0	0	0	0	0	0	0	0	0	0	0

F1 x F2

	focal	associated	any other with 2 NE
focal	87	100	0
associated	100	0	7.69

F1 x F3

	focal	associated	column 1	column 2	column 3	column 4	column 6	column 7	column 8	column 10	column 11	column 12	column 13	column 14	column 15
focal	87	100	0	0	0	0	0	0	0	0	0	0	0	0	0
associated	100	0	0	0	0	0	0	100	0	0	0	0	0	0	0

F2 x F3

	focal	associated	column 1	column 2	column 3	column 4	column 6	column 7	column 8	column 10	column 11	column 12	column 13	column 14	column 15
focal	87	100	0	0	0	0	0	0	0	0	0	0	0	0	0
associated	100	0	0	0	0	0	0	100	0	0	0	0	0	0	0
any other with 2 NE	0	100	15.38	15.38	15.38	15.38	15.38	7.69	15.38	15.38	15.38	15.38	15.38	15.38	15.38

Figure 16: Expected Payoffs under Asymmetric Frame Combinations BDV Game

F0 x F0	*any*
any	0.2

F1 x F1	*any black*	*any white*
any black	0.5	0
any white	0	0.33

F2 x F2	*any black*	*distinct white*	*any other white*
any black	0.5	0	0
distinct white	0	1	0
any other white	0	0	0.5

F3 x F3	*"b1"*	*"b2"*	*distinct white*	*"w1"*	*"w2"*
"b1"	0.5	0.5	0	0	0
"b2"	0.5	0.5	0	0	0
distinct white	0	0	1	0	0
"w1"	0	0	0	0.5	0.5
"w2"	0	0	0	0.5	0.5

Figure 17: Expected Payoffs under Symmetric Frame Combinations BG Game

F0 x F1	any black	any white
any	0.2	0.2

F0 x F2	any black	distinct white	any other white
any	0.2	0.2	0.2

F0 x F3	"b1"	"b2"	distinct white	"w1"	"w2"
any	0.2	0.2	0.2	0.2	0.2

F1 x F2	any black	distinct white	any other white
any black	0.5	0	0
any white	0	0.3	0.3

F1 x F3	"b1"	"b2"	distinct white	"w1"	"w2"
any black	0.5	0.5	0	0	0
any white	0	0	0.3	0.3	0.3

F2 x F3	"b1"	"b2"	distinct white	"w1"	"w2"
any black	0.5	0.5	0	0	0
distinct white	0	0	1	0	0
any other white	0	0	0	0.5	0.5

Figure 18: Expected Payoffs under Asymmetric Frame Combinations BG Game

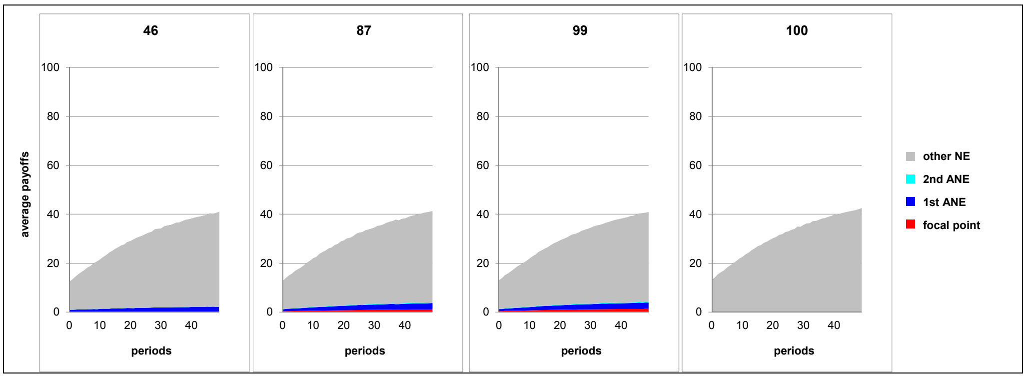


Figure 19: Average payoffs learning model without frames (fitting overall payoffs only)

References

Nabil I. Al-Najjar and Mallesh M. Pai. Coarse decision making and overfitting. *Journal of Economic Theory*, 150:467 – 486, 2014.

Larbi Alaoui and Antonio Penta. Endogenous depth of reasoning. *The Review of Economic Studies*, 83(4):1297–1333, 2015.

Federica Alberti, Robert Sugden, and Kei Tsutsui. Salience as an emergent property. *Journal of Economic Behavior & Organization*, 82(2):379–394, 2012.

Ayala Arad and Ariel Rubinstein. Multi-dimensional iterative reasoning in action: The case of the colonel blotto game. *Journal of Economic Behavior & Organization*, 84(2):571–585, 2012.

Ayala Arad and Ariel Rubinstein. Multi-dimensional reasoning in games: Framework, equilibrium and applications. *American Economic Journal: Microeconomics*, 2018.

Jasmina Arifovic and John Ledyard. A behavioral model for mechanism design: Individual evolutionary learning. *Journal of Economic Behavior & Organization*, 78(3):374–395, 2011.

Yaron Azrieli. Categorizing others in a large game. *Games and Economic Behavior*, 67(2):351–362, 2009.

Yaron Azrieli. Categorization and correlation in a random-matching game. *Journal of Mathematical Economics*, 46(3):303–310, 2010.

Michael Bacharach. Variable universe games. In Ken Binmore, Alan Kirman, and Piero Tani, editors, *Frontiers of Game Theory*, page 255. MIT Press, 1993.

Michael Bacharach and Michele Bernasconi. The variable frame theory of focal points: An experimental study. *Games and Economic Behavior*, 19(1):1–45, 1997.

Michael Bacharach and Dale O. Stahl. Variable-frame level-n theory. *Games and Economic Behavior*, 32(2):220–246, 2000.

Brent Berlin and Paul Kay. *Basic color terms: Their universality and evolution.* University of California Press, 1991.

Ken Binmore and Larry Samuelson. The evolution of focal points. *Games and Economic Behavior*, 55(1):21 – 42, 2006.

Andreas Blume and Uri Gneezy. Cognitive forward induction and coordination without common knowledge: An experimental study. *Games and Economic Behavior*, 68(2):488–511, 2010.

Tilman Börgers and Rajiv Sarin. Learning through reinforcement and replicator dynamics. *Journal of Economic Theory*, 77(1):1–14, 1997.

Antoni Bosch-Domènech and Nicolaas J. Vriend. On the role of non-equilibrium focal points as coordination devices. *Journal of Economic Behavior and Organization*, 94(0):52 – 67, 2013.

Robert R Bush and Frederick Mosteller. A mathematical model for simple learning. *Psychological review*, 58(5):313, 1951.

Colin F. Camerer, Teck-Hua Ho, and Juin-Kuan Chong. A cognitive hierarchy model of games. *Quarterly Journal of Economics*, 119(3):861–898, 2004.

André Casajus. Focal points in framed strategic forms. *Games and Economic Behavior*, 32(2):263–291, 2000.

André Casajus. *Focal Points in Framed Games (Lecture Notes in Economics and Mathematical Systems 499).* Springer-Verlag, 2001.

Miguel A. Costa-Gomes and Vincent P. Crawford. Cognition and behavior in two-person guessing games: An experimental study. *American Economic Review*, 96:1737–1768, 2006.

Jan Salomon Cramer. *Logit Models from Economics and Other Fields*. Cambridge University Press, 2003.

Vincent P. Crawford and Hans Haller. Learning how to cooperate: Optimal play in repeated coordination games. *Econometrica*, 58:571–595, 1990.

Vessela Daskalova and Nicolaas J. Vriend. Categorization and coordination. Technical report, Working Paper, 2019.

Ido Erev and Alvin E. Roth. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, 88:848–881, 1998.

Drew Fudenberg and David K Levine. *The theory of learning in games*, volume 2. MIT press, 1998.

David Gauthier. Coordination. *Dialogue: Canadian Philosophical Review/Revue canadienne de philosophie*, 14:195–221, 6 1975.

Robert Gibbons, Marco LiCalzi, and Massimo Warglien. What situation is this? Coarse cognition and behavior over a space of games. Working Papers 09, Department of Management, Universit Ca' Foscari Venezia, September 2017.

Sanjeev Goyal and Maarten Janssen. Can we rationally learn to coordinate? *Theory and Decision*, 40(1):29–49, 1996.

Veronika Grimm and Friederike Mengel. An experiment on learning in a multiple games environment. *Journal of Economic Theory*, 147(6):2220–2259, 2012.

Yuval Heller and Eyal Winter. Rule rationality. Mimeo, 2014.

Steffen Huck, Philippe Jehiel, and Tom Rutter. Feedback spillover and analogy-based expectations: A multi-game experiment. *Games and Economic Behavior*, 71(2):351–365, 2011.

Maarten C.W. Janssen. Rationalizing focal points. *Theory and Decision*, 50(2): 119–148, 2001.

Philippe Jehiel. Analogy-based expectation equilibrium. *Journal of Economic Theory*, 123(2):81–104, 2005.

Philippe Jehiel and Frdric Koessler. Revisiting games of incomplete information with analogy-based expectations. *Games and Economic Behavior*, 62(2):533 – 557, 2008.

Philippe Jehiel and Dov Samet. Valuation equilibrium. *Theoretical Economics*, 2(2):163–185, 2007.

John Maynard Keynes. *A Treatise On Probability.* Watchmaker Publishing, 2007.

Marco LiCalzi and Roland Mühlenbernd. Categorization and cooperation across games. *Games*, 10(1), 2019.

Daniel L. McFadden. Conditional logit analysis of qualitative choice behavior. In Paul Zarembka, editor, *Frontiers in Econometrics*, pages 105–142. Academic Press: New York, 1974.

Friederike Mengel. Learning across games. *Games and Economic Behavior*, 74 (2):601–619, 2012a.

Friederike Mengel. On the evolution of coarse categories. *Journal of Theoretical Biology*, 307:117–124, 2012b.

John Stuart Mill. *A System of Logic: Ratiocinative and Inductive*, volume 2. Harper, 1858.

Erik Mohlin. Optimal categorization. *Journal of Economic Theory*, 152:356–381, 2014.

Rosemarie Nagel. Unraveling in guessing games: An experimental study. *American Economic Review*, 85(5):1313–1326, 1995.

William S Neilson and Michael K Price. Focal strategies and behavioral backward induction in centipede games. Manuscript, University of Tennessee, Knoxville, 2011.

Willard Van Orman Quine. Natural kinds. In *Essays in honor of Carl G. Hempel*, pages 5–23. Springer, 1969.

Julian Romero and Yaroslav Rosokha. A model of adaptive reinforcement learning. Technical report, SSRN Working Paper, 2019.

Alvin E. Roth and Ido Erev. Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, 8(1):164–212, 1995.

Thomas C Schelling. *The strategy of conflict*. Harvard university press, 1980.

Dale O. Stahl and Paul W. Wilson. Experimental evidence on players' models of other players. *Journal of Economic Behavior and Organization*, 25(3):309–327, 1994.

Dale O. Stahl and Paul W. Wilson. On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior*, 10(1):218–254, 1995.

Robert Sugden. A theory of focal points. *Economic Journal*, 105:533–550, 1995.

Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT Press, 2018.

H Peyton Young. The evolution of conventions. *Econometrica*, pages 57–84, 1993.