DISCUSSION PAPER SERIES

DP14088

NAIVETE AND SOPHISTICATION IN INITIAL AND REPEATED PLAY IN GAMES

Nagore Iriberri and Bernardo Garcia-Pola

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

NAIVETE AND SOPHISTICATION IN INITIAL AND REPEATED PLAY IN GAMES

Nagore Iriberri and Bernardo Garcia-Pola

Discussion Paper DP14088
Published 31 October 2019
Submitted 29 October 2019

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

NAIVETE AND SOPHISTICATION IN INITIAL AND REPEATED PLAY IN GAMES

Abstract

Naive, non-equilibrium, behavioral rules, compared to more sophisticated equilibrium theory, are often better in describing individuals' initial play in games. Additionally, in repeated play in games, when individuals have the oppor- tunity to learn about their opponents' past behavior, learning models of different sophistication levels are successful in explaining how individuals modify their be- havior in response to feedback. How do subjects following different behavioral rules in initial play modify their behavior after learning about past behavior? This study links both initial and repeated play in games, analyzing elicited be- havior in 3x3 normal-form games using a within-subject laboratory design. We classify individuals into different behavioral rules in both initial and repeated play and test whether and/or how naivete and sophistication in initial play cor- relates with naivete and sophistication in repeated play. We find no evidence for a correlation between naivete and sophistication in initial and repeated play.

JEL Classification: C70, C91, C92

Keywords: Naivete, Sophistication, initial play, repeated play, level-k thinking, adaptive and sophisticated learning, mixture-of-types estimation

Nagore Iriberri - nagore.iriberri@gmail.com
University of the Basque Country and CEPR

Bernardo Garcia-Pola - bernardo.garcia@ehu.eus
University of New South Wales

Naivete and Sophistication in Initial and Repeated Play in Games*

Bernardo García-Pola[†] Nagore Iriberri[†‡]

October 29, 2019

Abstract

Naive, non-equilibrium, behavioral rules, compared to more sophisticated equilibrium theory, are often better in describing individuals' initial play in games. Additionally, in repeated play in games, when individuals have the opportunity to learn about their opponents' past behavior, learning models of different sophistication levels are successful in explaining how individuals modify their behavior in response to feedback. How do subjects following different behavioral rules in initial play modify their behavior after learning about past behavior? This study links both initial and repeated play in games, analyzing elicited behavior in 3×3 normal-form games using a within-subject laboratory design. We classify individuals into different behavioral rules in both initial and repeated play and test whether and/or how naivete and sophistication in initial play correlates with naivete and sophistication in repeated play. We find no evidence for a correlation between naivete and sophistication in initial and repeated play.

Keywords: Naivete, sophistication, initial play, repeated play, level-k thinking, adaptive and sophisticated learning, mixture-of-types estimation

*Project supported by a 2018 Leonardo Grant for Researchers and Cultural Creators, BBVA Foundation. The Foundation accepts no responsibility for the opinions, statements and contents included in the project and/or the results thereof, which are entirely the responsibility of the authors. We would also like to thank Pedro Rey-Biel and seminar attendees of various presentations for their useful comments.

†AGORA Center, Dept. of Economics, UNSW Business School, Sydney, NSW 2052, Australia (bernardo.garciapola@gmail.com).

‡University of the Basque Country, EHU-UPV, and IKERBASQUE, Basque Foundation for Research (nagore.iriberri@gmail.com)

1 Introduction

Nash equilibrium has been and still is the benchmark solution concept in game theory for predicting individual behavior in strategic environments. Since economics adopted the use of laboratory experiments, hundreds of experimental studies have tested whether individual behavior complies with the predictions of Nash equilibrium theory. These studies have shown that equilibrium theory has clear limitations in regard to its ability to describe how people behave in strategic environments. In reaction to the ample experimental evidence, the important contributions of behavioral economics include models of bounded rationality that improve our understanding of how people actually behave in two different domains. First, when individuals make decisions for the first time, with no previous experience or opportunity to learn, a scenario called initial play, naive, non-equilibrium, behavioral rules often exceed equilibrium theory in their ability to describe individual behavior (see for example, Goeree and Holt, 2001, and Crawford et al., 2013). Second, given that people often do not start playing Nash equilibrium strategy, bounded rationality models have been applied to repeated play to understand how people modify their behavior when provided with feedback about past behavior, that is, models that explain how individuals learn over time (see for example, Sobel, 2000).

Does behavior in initial responses relate in any way to behavior in repeated play? This is the central question of this paper.

When studying initial play, models that explain how individual begin playing games differ in the naivete or sophistication assumed with respect to individual thinking in strategic environments. We can order the behavioral rules in initial play from the most naive to the most sophisticated. The most naive behavioral rules include processes that require no *strategic* thinking, such that strategic settings are considered to be isomorphic to pure decision making settings. For example, maxmax and maxmin behavioral rules fall into this category because maximizing over possible outcomes, or maximizing over minimum possible outcomes requires no ability to predict an opponent's behavior. Level-k thinking models, which have been shown to be successful in explaining initial behavior in different settings (Stahl and Wilson, 94, 95; Nagel, 95; Costa-Gomes et al., 2001, Camerer et al., 2004), illustrate quite well different levels of sophistication. Level-*1* behavioral type calculates the expected payoff associated with each of the available strategies, assuming each of the opponent's actions is equally likely, and takes the strategy with the highest expected payoff. In the spirit of this interpretation, we also consider level-*1* to be a naive behavioral model.[1] More sophisticated behavioral rules assume that individuals are best responding to some type of opponent behavior but assume that opponents follow a more naive behavioral rule than themselves. Level-*2* and level-*3* represent an increasing sophistication assumed about the opponent's action, as level-*2* believes the opponent is behaving as a naive level-*1* and best responds to those beliefs, while level-*3* assumes the opponent behaves as a level-*2* and best responds to those beliefs. Finally, among the most sophisticated behavioral rules is the Nash equilibrium, which considers not only common knowledge of rationality but rational expectations about beliefs.

In studies focused on repeated play, models that explain how individuals modify their behavior in response to feedback on (own and opponent's) past behavior also differ in the naivete or sophistication assumed with respect to individual behavior in a strategic environment. Learning models can also be ordered according to their sophistication level, from the most naive to the most sophisticated, in a hierarchical manner. The most naive learning model includes a behavioral rule that ignores any type of feedback and repeats the same strategy as used in the past. We refer to this as the *No-Change* behavioral rule in repeated play. Adaptive learning models assume that individuals modify their behavior in response to feedback, i.e. best responding to an opponent's past behavior (illustrated best by fictitious play, as in Fudenberg and Levine, 98a and 98b). Note that adaptive learners assume that opponents indeed follow a *No-Change* type, as they assume that opponents will repeat the same strategy as used in the past; therefore adaptive learners will best respond to their past strategy. Finally, more sophisticated learning models assume that opponents indeed learn through an adaptive learning model and accordingly best respond to this (see for example, Milgrom and Roberts, 91; Selten, 91; Conslik, 93a and 93b; Nagel, 95; Camerer et al., 2002, and Stahl, 2003).

Somewhat surprisingly, the literature on learning models (i.e., Cheung and Friedman, 97; Erev and Roth, 98, Fudenberg and Levine, 98a and 98b, and Camerer and Ho,

[1]The cognitive hierarchy model (Camerer et al., 2004) assumes that level-k players best respond to combinations of existing lower levels. However, both level-k thinking and cognitive hierarchy models coincide in terms of the level-*1*'s predictions.

99) and the literature on models to explain initial behavior have followed parallel paths (summarized in Crawford et al., 2013).[2] On the one hand, when studying learning over time, initial play has been treated as a “black box”, an exogenous factor used only to initialize learning models, for example estimating initial “attractions” associated with each of the particular strategies or, alternatively, simply assuming that initial “attractions” are the same across different strategies. On the other hand, models that aim to explain initial behavior have used mostly experimental designs that provide no feedback from game to game, precisely to suppress any opportunity to learn. Such models have been silent on explaining learning over time.

However, it appears to be natural that some type of relation exists between behavior in initial and repeated play. As Costa-Gomes and Crawford (2006) note, modeling initial responses more precisely could yield insights into cognition that elucidate other forms of strategic behavior, such as learning and distinguishing between different levels of sophistication in rules and therefore, influencing implications for equilibrium selection and convergence. However, similar consistencies that seemed a priori intuitive have been empirically rejected (Costa-Gomes and Weizsacker, 2008; Knoepfle et al., 2009). The question of whether the behavior in these two contexts is related arises not only as a natural question but as an important one. If such behavior is related, observing the initial behavior of an individual would be informative of how her behavior will change and vice versa. Furthermore, this relation would allow a unified framework of behavior in games that incorporates both initial and repeated play (see, for example, Ho et al., 2015). If such behavior is not related, such that we observe very different levels of sophistication when the same individual faces a situation for the first time and in repeated play, characteristics that we sometimes measure as inherent to an individual, like cognitive ability may, be more context-dependent than previously believed.

We therefore study fundamental questions for/when proposing a unified framework for studying initial and repeated play in games: How do naivete and sophistication in initial play relate to naivete and sophistication in repeated play? Is a naive player in initial responses, compared to a more sophisticated player, more likely to learn through

[2]There are few exceptions, as some models have been used to explain both initial behavior and learning behavior over time, such as quantal response equilibrium by McKelvey and Palfrey (95), simply estimating different noise levels or lambda-s for behavior in different stages.

a naive learning model in repeated play? We propose a laboratory experiment and a mixture-of-types model econometric estimation to address these inherently empirical question.

Subjects in our experiment go over 14 different 3×3 games (actually 7 games, where subjects play both as row and column players) two times in two different stages of the experiment. In the first stage, subjects receive no feedback from game to game, with the objective of eliciting their initial play (with no opportunity to learn or obtain experience from game to game). Based on the subjects' profiles of 14 decisions, we classify each subject as following one of multiple behavioral rules. This exercise is similar to those pioneered by Sthal and Wilson (94, 95), and later used by for example Costa-Gomes et al. (2001), Costa-Gomes and Crawford (2006), Rey-Biel (2009) and García-Pola et al., (2016). In the second stage of the experiment, subjects repeat the same 14 games, but this time, in each of the games they receive feedback on what they did in the first stage, as well as what their *current* opponent did in the first stage. Using subjects' profiles of 14 decisions and observed feedback on their own and current opponent's past strategies, we classify each subject as following one of multiple behavioral rules in repeated play. Note that this elicitation and identification of learning rules is different from studies that attempt to identify the ability of different learning rules to explain behavioral data (see, for example Erev and Roth, 98; Camerer and Ho, 99; Feltovich, 2000, and more recently, Kovarik et al., 2018). In our setting, for a particular game, subjects can learn about an opponent's past actions just once, but we elicit how subjects learn from 14 different games or decisions based on opponents' past action in those 14 different games. In other words, we elicit subjects' learning rules using multiple different games in a way that subjects themselves cannot evaluate how successful their learning model is, which we refer to as "initial model of learning". As this study is, as far as we know, the first empirical exercise that connects initial and repeated play, we designed games with the purpose of allowing for the highest separation among different behavioral rules in both initial and repeated play. The separation is the cornerstone for the use of a mixture-of-types model when identifying and classifying subjects into different behavioral rules, both in initial and repeated play. Finally, the within-subject design allows us to construct contingency tables to test whether naivete and sophistication in initial play are correlated with naivete and sophistication in repeated play.

We find no evidence for a correlation between the naivete and sophistication in the initial and repeated play. Regarding initial behavior, consistent with previous findings, we find that the majority of subjects, 60% of them, use a naive behavioral rule. The second most frequent rule is a more sophisticated behavioral rule, level-*2*, used by 36% of the subjects. Additionally, consistent with previous findings, few Nash equilibrium players are found among the subject population. Furthermore, when identifying the behavioral rules that describe repeated play, the majority of subjects, 57%, also show the most naive behavioral rule of ignoring their opponent's past action, followed by adaptive learners, 28%, who consistently best respond to the opponent's past strategy. Sophisticated learning models are also rarely used. Most importantly, and surprisingly, when naivete and sophistication are compared between initial and repeated play, which is the central question of our study, subjects using a naive behavioral rule and more sophisticated players (level-2) in initial play show a similar likelihood of using the most naive (*No-Change* type) and adaptive learning model in repeated play. This finding shows little support for any correlation between naivete and sophistication across initial and repeated play.

The rest of this paper is organized as follows. Section 2 describes the experimental procedures and design in detail. Section 3 presents the results, divided into the identification and classification of subjects according to their initial play, identification and classification of subjects according to their repeated play, and the correlation between naivete and sophistication across the two settings. Section 4 includes two important robustness tests of the potential misspecification in the identification and classification of behavioral rules. Finally, section 5 concludes the paper.

2 Experimental Procedures and Design

2.1 Procedures

A total of 198 subjects who participated in the experiment were recruited using ORSEE system (Greiner, 2015). The sessions were conducted via computer using z-Tree software (Fischbacher, 2007). In April and May of 2019, two sessions with a total of 78 subjects took place in the Laboratory of Experimental Analysis (Bilbao Labean; http://www.bilbaolabean.com) at the University of the Basque Country. We conducted two additional sessions with the remaining 120 subjects in the Laboratory of Experimental Economics (LEE, http://lee.uji.es) at the University Jaume I of Castellón.

Subjects were told that the experiment consisted of two different parts and that payments would depend on both luck and their own and other subjects' decisions. Before each part, subjects were given detailed instructions explaining the task involved, including examples of games, how they could make decisions and how they were going to be matched and paid. Subjects were allowed to ask any question they may have during the instructions. At the end of the instructions, subjects were asked a few questions to guarantee that they understood the instructions regarding each part. A translated version of the instructions can be found in Appendix C.

All subjects played the same seven 3×3 normal-form two-player games in the same order, first as the row player and then as the column player, playing a total of 14 games in each part. We did not inform subjects that they played the same games as different roles, and we showed all the games to all subjects from the perspective of row players. Subjects were randomly matched in a way that, within each part of the experiment, they were paired with a different opponent in each of the 14 games. In the first part of the experiment, subjects received no feedback from game to game to elicit initial play in the 14 games. In the second part, subjects repeated the same 14 games in the same order but now they were given feedback about their own past strategy and their current opponent's past strategy in the first part of the experiment. The fact that subjects will be provided with feedback in the second part was public knowledge. An example of how games in both parts and feedback in the second part were displayed in the experiment can be found in the instructions in Appendix C.

When all subjects had submitted their choices in the two parts, for each subject, the computer randomly chose two games from any of the two parts for payment. Thus, each subject could be paid for different games. Before being paid, subjects completed a non-incentivized questionnaire regarding demographic data (gender, age, field of study, nationality), risk preferences following Eckel and Grossman (2002), and cognitive ability using cognitive reflection test. Descriptive statistics of all these variables can be found in Appendix Table A1. The subject pool shows the typical characteristics of undergraduate students who come mostly from Economics and Business degrees, with a slightly higher presence of females, given most are pursuing a degree in social sciences. We also requested free-format explanations for their choices and expected choices of others in each of the parts of the experiment. We did not include these data in the analysis, but we did informally assess the consistency between subjects' explanations of what they did and the rule we estimated using their elicited actions and frequently observed a clear coincidence between the two. For a work that attempts to relate subjects' free-format explanations of their actual actions and their actions, see Brañas et al. (2011). Finally, we paid subjects privately according to the two games selected plus a 3 Euro show-up fee. The average payment was 15.76 Euros, with a standard deviation of 4.90. The entire experiment lasted one hour and a half including the reading of instructions and payment.

2.2 Design of Games

We designed seven 3×3 normal-form games, as shown in Figure 1. The actual order in which the games were presented to subjects was: G1, G2,... until G7 as row players, to which we will refer as G11, G21, and so on until G71, and G1, G2,... until G7 as column players, to which we will refer as G12, G22, and so on until G72. As noted in the previous section, all subjects were presented the games as if they were row players, that is, we transposed the games when the subjects were playing as column players. We chose the particular sequence, first as row players and then as column players, to avoid subjects realizing that they were making choices in the same games.

We chose 3×3 normal-form games because such games allow for ample separation between the predictions of different behavioral rules. Note that with 14 3×3, there are 4,782,969 possible strategy profiles, while with 2×2 games, we would have *only* 16,384 possible strategy profiles. Therefore, having 3×3 games substantially increases the a priori possibility of separation among the predictions of different behavioral rules. Additionally, we chose 3×3 games, instead of, for example, 4×4 games, to ensure that the number of strategies was relatively small such that it was easy to handle by subjects, which facilitated the explanation of the instructions. Finally, we designed our own games instead of using games from other studies, because we aimed to have high separation between different behavioral rules both in initial and repeated play, which was not the aim of any of the previous studies.

Figure 1: Experimental Games

G1

4 , 20	20 , 12	18 , 2
6 , 8	8 , 14	22 , 16
18 , 14	14 , 6	2 , 18

G2

6 , 18	22 , 4	4 , 16
20 , 6	2 , 24	16 , 4
12 , 12	2 , 6	18 , 22

G3

4 , 20	12 , 16	16 , 4
18 , 2	20 , 12	2 , 8
22 , 18	2 , 2	22 , 10

G4

10 , 18	20 , 16	4 , 6
12 , 10	14 , 22	2 , 12
6 , 4	18 , 4	16 , 18

G5

8 , 16	16 , 14	20 , 12
16 , 8	18 , 12	4 , 4
14 , 6	16 , 4	2 , 20

G6

14 , 16	2 , 20	12 , 22
6 , 18	20 , 4	10 , 6
22 , 4	14 , 18	4 , 10

G7

4 , 20	22 , 14	18 , 4
6 , 6	8 , 12	20 , 14
18 , 16	14 , 8	4 , 18

2.3 Assessing the Design: Behavioral Rules' Predictions and Separation across Games

The 3×3 normal-form games were carefully designed with the aim of having the largest separation between the predictions of different behavioral rules. We now explain in detail which behavioral rule we consider in each of the two parts of the experiment and describe the predicted strategies across the 14 games. We finish this section by showing the actual separation between the predicted behavior.

As noted, there are 14 different games, eachwith 3 available strategies, so there are 4,782,969 possible strategy profiles. To understand how individuals make decisions in these 14 games, we consider a handful of behavioral rules, taken from the literature both in initial and repeated play.

We consider 8 behavioral types when attempting to explain subjects' initial play in games. We take the leading behavioral models from in the literature (Stahl and Wilson, 94 and 95, Nagel, 95, Costa-Gomes et al. 2001, Costa-Gomes and Crawford, 2006, and García-Pola et al., 2016, among others). We distinguish between non-strategic behavioral rules, i.e., those that do not need to anticipate the opponent's strategy, and strategic behavioral rules, i.e., those that do need to anticipate the opponent's strategy. Among the non-strategic behavioral rules we consider: altruistic or social welfare maximizer (*A*), inequity-averse (*IA*), maxmax or optimistic (*MaxMax*), maxmin or pessimistic (*MaxMin*) and level-*1* type (*L1*). Among the strategic behavioral rules we consider level-*2* (*L2*t), level-*3* (*L3* in short) and Nash equilibrium play (*NE*). For a clearer exposition, we now present the rules from lowest to highest sophistication level in terms of strategic thinking.

The altruistic type simply sums own and opponent's payoffs and takes the strategy that leads to the highest sum of payoffs. The IA type takes the absolute value of the difference between the own and opponent's payoff and takes the strategy that leads to the lowest difference. The optimistic type follows the strategy that leads to the maximum possible own payoff, while the pessimistic type follows the strategy that maximizes the minimum possible own payoff. The level-*1* type sums own payoffs across columns (opponent's three possible strategies) and takes the strategy that yields the maximum expected payoff. Level-*1* can be considered to be at the edge between non-strategic and strategic behavioral rules, as it can also be interpreted as the best response to uniform play by the opponent. Level-*2* expects the opponent to behave as level-*1* type and best responds to those beliefs. Level-*3*, similarly, expects the opponent to behave as level-*2* type and best responds to those beliefs. Finally, NE play calculates the mutual best response required by equilibrium thinking. The top panel of Table 1 shows the predictions of each of these behavioral rules in the 3×3 normal-form games in Figure 1.

We consider 4 behavioral types when trying to explain subjects' repeated play with feedback in games. We also take the leading behavioral models from the literature (Fudenberg and Levine, 98a and 98b; Nagel, 95; Camerer et al., 02; Stahl, 03) and present them according to their sophistication level as well. The simplest or most naive behavioral rule one can consider is the no-change type (*No-Change*), which simply ignores the provided feedback and mimics the behavior taken in the first part of the experiment. Adaptive learning behavior (*Adaptive*) assumes that individuals best respond and that they try to guess what their opponent will do (similar to any belief-based learning model, as in Fudenberg and Levine, 1998). In our setting, as subjects are provided with the opponent's past strategy, adaptive learning assumes that the opponent will repeat her/his past strategy, so opponents are expected to follow a *No-Change* type and therefore best responds to such behavior.[3] Sophisticated learning

[3]Notice that in our repeated play setting, given that subjects are never provided with how successful

(*Sophisticated*) goes one step further and considers that the opponent follows adaptive learning behavior. As such, the sophisticated learning rule uses own past behavior, and calculates the corresponding adaptive learning behavior (i.e., best response to own past behavior) and then best responds to those beliefs regarding the opponent's expected behavior. Finally, we also consider one more round of sophistication in repeated behavior (*Sophisticated 2*). *Sophisticated 2* learning type assumes that the opponent is follows sophisticated learning behavior (best response to own behavior as an adaptive learner) and best responds to those beliefs. Note that all these behavioral types not only require a particular game to make predictions but also need own and/or opponents' past behavior, so they are dependent on observed past behavior. The bottom panel of Table 1, therefore, does not show the actual predicted strategies, but in general, the calculation of a particular behavioral rule requires repeated play with feedback.

Table 1: Predicted Strategies by Different Behavioral Rules

	G11	*G12*	*G21*	*G22*	*G31*	*G32*	*G41*	*G42*	*G51*	*G52*	*G61*	*G62*	*G71*	*G72*
Initial Play														
A	2	3	3	1,2	1	1	1	3	3	1	2	3	3	2
IA	2	3	3	1,2,3	2	1	2	1	1	2	1,3	3	1	1
MaxMax	2	1	3	1	1	3	1	1	2	3	2	3	3	1
MaxMin	2	1	1	3	1	2	2	1	1	3	3	1	3	2
L1	1	2	2	3	1	3	1	1	3	1	2	3	2	1
L2	3	3	3	1	1	2	3	1	2	2	3	1	2	1
L3	3	1	2	3	2	2	3	3	3	3	1	1	1	3
NE	3	1	2	3	2	2	3	3	3	3	1	1	1	3
Repeated Play														
No-Change	"Same strategy as in the first part"													
Adaptive	"Best response to (opponent's past strategy)"													
Sophisticated	"Best response to (opponent's best response to (own past strategy))"													
Sophisticated 2	"Best response to (opponent's best response to (best response to (opponent's past strategy)))"													

Notes: The table reports the strategies predicted by the models in initial play (top panel) and the models in repeated play (bottom panel). 1, 2 and 3 refer to first, second and third strategy, respectively. In a few instances, a behavioral rule is indifferent between multiple strategies, so we assume the behavioral rule will predict any of those strategies with equal probability.

Finally, Table 2 shows the separation between different behavioral rules, in both initial play (panel A) and repeated play with feedback (panel B). The values in the

their past strategy in the first stage was, reinforcement learning (Erev and Roth, 98) cannot be directly assessed. However, with a more flexible interpretation and assuming that subjects evaluate their past strategy with the current opponent's past strategy, reinforcement and adaptive learning models would predict the same strategy.

table represent the proportion of games in which the predictions of two behavioral rules (the one in the row and the one in the column) are separated. The numbers can take any value between 0 (no separation at all, such that two behavioral rules predict exactly the same strategy in each of the 14 games) and 1 (full separation, such that two behavioral rules predict a different strategy in each of the 14 games).

Table 2: Separation of Different Behavioral Rules

Panel A: Initial Play

	A	*IA*	*MaxMax*	*MaxMin*	*L1*	*L2*	*L3*	*NE*
A	0.00							
IA	0.60	0.00						
MaxMax	0.46	0.62	0.00					
MaxMin	0.71	0.65	0.57	0.00				
L1	0.57	0.76	0.50	0.79	0.00			
L2	0.75	0.58	0.57	0.64	0.71	0.00		
L3	0.86	0.80	0.86	0.64	0.79	0.71	0.00	
NE	0.57	0.51	0.86	0.64	0.79	0.79	0.57	0.00

Panel B: Repeated Play with Feedback

	No Change	*Adaptive*	*Sophisticated*	*Sophisticated 2*
No Change	0.00			
Adaptive	0.65	0.00		
Sophisticated	0.71	0.60	0.00	
Sophisticated 2	0.71	0.60	0.47	0.00

Notes: The table reports the proportion of strategies across all 14 games in which the different behavioral models predict different strategies. The minimum possible separation value is 0, when the two models prescribe the same strategy in all 14 games, and the maximum possible separation value is 1, when two models predict a different strategy in each of the 14 games.

The separation values for the initial play range between 0.46 and 0.86, which shows that each pair of behavioral rules is separated in at least 6 of 14 games, and as many as 12 out 14 games. Regarding the separation values in repeated play with feedback, we could not calculate these values ex-ante, as they depended on the particular observed past behavior of subjects.[4] The values in panel B are therefore based on the actual

[4]We could indeed use, as we actually did, the accumulated evidence from past studies (see Crawford et al., 2013, for example) that found that approximately half of the subject population showed non-

observed behavior in the first part of the experiment. The values range between 0.47 and 0.71, which indicates that two behavioral rules for repeated play with feedback are separated in at least 6 of 14 games, and, as many as 10 of 14 games. To put these separation values into perspective, note that any behavioral rule will have a separation value from a randomly generated behavioral rule of 0.67, that is, given that there are 14 games, each with 3 possible strategies, a randomly generated behavioral rule will coincide with an existing behavioral rule in 4.6 games. We therefore conclude that the goal of attaining a large separation between the considered behavioral rules is achieved.

3 Results

3.1 Descriptive Overview

We begin by considering the mean behavior in both initial and repeated play, which represents how individuals start playing in strategic environments with no feedback (first part) and how individuals react to both their own and current opponent's past behavior (second part).

Figure 2 shows the results for the first (panel A) and second parts (panel B) of the experiment. Clearly, individual behavior is different from random play in both initial and repeated play; otherwise, we would observe that in each game, each of the 3 strategies is played with 1/3 probability (p-values less than 0.001 for both the first and second parts, using a chi-square test against a uniform distribution). Additionally, the mean behavior does not differ significantly between the first and second parts of the experiment, as we cannot reject that the behavior in both scenarios comes from the same distribution (p-value of 0.84 from the two sample chi-square test that two data samples come from the same distribution), which already suggests that many subjects ignore the feedback and follow the same strategy as in the first part. The key task in the next two subsections is to identify the relevant behavioral types that are able to reproduce the behavior in both parts of the experiment.

strategic behavior and a smaller proportion more sophisticated behavioral rules such as *L2* and *L3* with a minority of subjects following Nash equilibrium strategy.

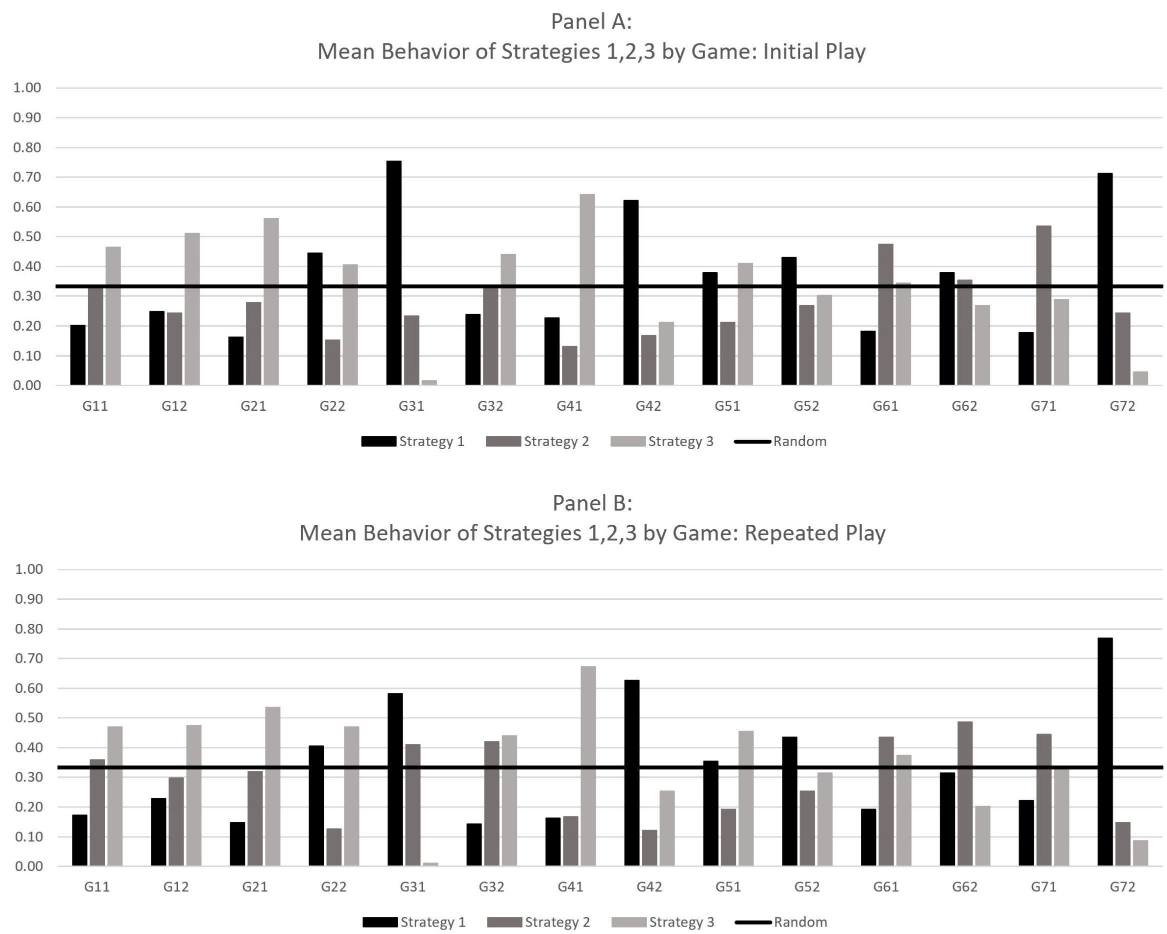


Figure 2: Mean Behavior in Initial Play and Repeated Play with Feedback

3.2 Naivete and Sophistication in Initial Play: Type Identification

Using individual data on revealed choices by 198 subjects in 14 different games in the first part of the experiment, we proceed to identify the behavioral type of each subject in initial play. Using a mixture-of-types model with uniform errors, we identify and classify each of the 198 subjects into a behavioral type. The maximum likelihood function is estimated subject by subject. Please see Appendix B for a general description of the maximum likelihood function used to estimate behavioral types and for a particular derivation of the maximum likelihood function for estimating the behavioral types in initial play.

Table 3 shows the estimation results. We allow for different errors or alternatively perfect guesses, from 7 to 11 perfect guesses. Note that, by chance, if individual play was random, any behavioral type that predicts a particular strategy profile, would make 4.6 perfect guesses. Therefore, using this value as a benchmark, we consider both less and more stringent identification criteria: no constraints, at least 7 perfect guesses (50% improvement over random), 9 (93% improvement over random) and 11 perfect guesses (139% improvement over random). As expected, a trade-off exists between the number of perfect guesses required for identification and the number of subjects we can properly identify. Nevertheless, remarkably, when imposing 9 perfect guesses (out of 14), which is a high threshold (93% improvement over random), we can identify 93 subjects.

Table 3: Behavioral Type Identification for Initial Play

	Minimum Number of Perfect Guesses			
Model	No Constraints	7	9	11
	(1)	(2)	(3)	(4)
Non-strategic	0.60	0.58	0.53	0.62
A	0.15	0.13	0.10	0.14
IA	0.07	0.06	0.01	0.00
$MaxMax$	0.11	0.11	0.09	0.05
$MaxMin$	0.12	0.11	0.11	0.10
L1	0.16	0.17	0.22	0.33
L2	0.36	0.37	0.47	0.38
L3	0.02	0.02	0.00	0.00
NE	0.02	0.02	0.01	0.00
No. of Subjects	198	186	93	21

Notes: The table displays the population frequencies estimated to be consistent with each of the behavioral rules listed in Model column for different numbers of perfect guesses, from all subjects (column 1) to subjects with 7, 9 and 11 (column 4) perfect guesses.

As observed in Table 3, focusing on the overall population, in column 1, 60% of the subjects follow a non-strategic behavioral rule, followed by *L2* (36%), and only a minority of subjects (4%) are identified as sophisticated *L3* and *NE*. Among the non-strategic behavioral types, *L1* and *A* explain most of the behavior, followed by pessimistic and optimistic behavioral rules. These results are roughly consistent with existing findings, although we find lower frequencies for *L1* and higher frequencies for *L2*. Furthermore, these conclusions do not change if we move across different columns (criteria over the required perfect guesses). Only when we impose a number of correct guesses of 11, for which we can only identify 21 subjects, we find considerably more

L1 individuals in detriment of the optimistic types. However, the overall conclusions remain unchanged: we still find that approximately 62% of the subject population is identified to follow a non-strategic behavioral rule, followed by *L2* (38%). We cannot reject that the type distribution of the subjects depends on the constraints imposed regarding the number of perfect guesses (p-value of 0.12 for the chi-square test), so the estimation results are robust to the criteria on the perfect guesses.

3.3 Naivete and Sophistication in Repeated Play with Feedback: Type Identification

Using individual data on revealed choices by 198 subjects in 14 different games in the second part of the experiment, we proceed to identify the behavioral type of each subject in repeated play. Using a mixture-of-types model with uniform errors we identify and classify each of the 198 subjects into a behavioral type. The maximum likelihood function is estimated subject by subject. Please see Appendix B for a general description of the maximum likelihood function used to estimate behavioral types and for a particular derivation of the maximum likelihood function for estimating the behavioral types in repeated play.

Table 4 shows the estimation results. As in the first part of the experiment, we allow for different errors or alternatively perfect guesses, from 7 to 11 perfect guesses. Note that, by chance, if individual play was random, any behavioral type that predicts a particular strategy profile, would make 4.6 perfect guesses. Therefore, using this value as a benchmark, we allow for less stringent to more stringent identification of behavioral types: no constraints, at least 7 perfect guesses (50% improvement over random), 9 (93% improvement over random) and 11 perfect guesses (139% improvement over random). Again, a trade-off exists between the number of required perfect guesses for identification and the number of subjects we can properly identify. The number of subjects we can cleanly identify is better than that in the first part. When we impose the criterion of 9 perfect guesses, we now identify 144 subjects (73% of the subject population).

The behavior of more than half of the subjects is best explained by the *No-Change* type, which reflects that the majority of subjects ignore the opponent's past behavior and simply repeat their own past behavior. We found such a high frequency of simply

Table 4: Behavioral Type Identification for Repeated Play

Model	Minimum Number of Perfect Guesses			
	No Constraints	7	9	11
	(1)	(2)	(3)	(4)
No-Change	0.57	0.56	0.63	0.67
Adaptive	0.28	0.28	0.26	0.29
Sophisticated	0.10	0.10	0.08	0.03
Sophisticated 2	0.06	0.06	0.03	0.01
No. of Subjects	198	188	144	70

Notes: The table displays the population frequencies estimated to be consistent with each of the behavioral rules listed in Model column for different number of perfect guesses, from all subjects (column 1) to subjects with 7, 9 and 11 (column 4) perfect guesses.

repeating play ignoring the opponent's past behavior to be surprising. The second most common behavior is adaptive behavior, followed by 28% of the subjects, that is, those who best respond to opponent's past behavior. Finally, very few subjects show sophisticated learning behavior. Consistent with the previous findings, it is reassuring that these conclusions do not change as we move across different columns. If anything, when the highest threshold of 11 perfect guesses is imposed, the frequency of *No-Change* increases by 10 percentage points to the detriment of *Sophisticated* the learning model. As before, we cannot reject that the type distribution of the subjects depends on the constraints imposed regarding the number of perfect guesses (p-value of 0.13 for the chi-square test), so the results are robust to the criteria on perfect guesses.

3.4 Correlation between Naivete and Sophistication in Initial and Repeated Play

We now study the central question of the paper, the correlation between the type identification in initial and repeated play, exploiting the fact that all the subjects participated in the same two parts of the experiment. We use a contingency table, where the rows present the behavioral rules in initial play and the columns present the behavioral rules in repeated play. Therefore, a particular cell in the contingency table shows the proportion of subjects identified as following the behavioral rule in that particular row in initial play who also follow the behavioral rule in that particular column in repeated play. The frequencies across the columns sum to 1 in each row. A positive correlation would show a higher frequency of a naive, non-strategic, behavioral rule in initial play to be using a *No-Change* or less sophisticated rules in repeated play than level-*2* subjects, who would show a higher frequency of learning as adaptive or sophisticated learners. A no correlation result would show independence in the distributions across different rows. A negative correlation would show that a naive behavioral rule in initial play is using a more sophisticated learning model than a more sophisticated rule in initial play.

Table 5: Contingency Table

Panel A: No constraints

	Second Part Model				
First Part Model	*No-change*	*Adaptive*	*Sophisticated*	*Sophisticated 2*	No. of Subjects
Non-strategic	0.53	0.30	0.11	0.06	119
A	0.55	0.28	0.14	0.03	29
IA	0.46	0.23	0.15	0.15	13
$MaxMax$	0.38	0.57	0.00	0.05	21
$MaxMin$	0.67	0.25	0.04	0.04	24
L1	0.53	0.22	0.19	0.06	32
L2	0.63	0.25	0.07	0.06	72
L3	0.50	0.25	0.25	0.00	4
NE	0.67	0.00	0.33	0.00	3
No. of Subjects	112	55	20	11	198

Panel B: Minimum of 9 correct guesses in each part

	Second Part Model				
First Part Model	*No-change*	*Adaptive*	*Sophisticated*	*Sophisticated 2*	No. of Subjects
Non-strategic	0.69	0.28	0.03	0.00	32
A	0.71	0.29	0.00	0.00	7
IA	-	-	-	-	0
$MaxMax$	0.50	0.50	0.00	0.00	4
$MaxMin$	0.88	0.13	0.00	0.00	8
L1	0.62	0.31	0.08	0.00	13
L2	0.69	0.23	0.05	0.03	39
L3	-	-	-	-	0
NE	-	-	-	-	0
No. of Subjects	49	18	3	1	71

Notes: The table shows the proportion of subjects identified as each of the behavioral rules of the first part that were also classified as each of the behavioral rules of the second part.

As observed in panel A of Table 5, for all 198 subjects, we see little evidence of correlation between the naivete and sophistication in initial and repeated play. Almost any behavioral type in initial play is equally likely to fall into the *No-Change* or *Adaptive* behavioral rules for repeated play. Indeed almost half of the subjects in the population fall into the most naive *No-Change* behavioral rule in repeated play, followed by the adaptive learning model (between 25 and 30% of the subjects) and a minority follow a more sophisticated behavioral rule, independent of the naivete and sophistication shown in initial play. Panel B shows the equivalent results for a reduced number of subjects when we impose the criterion of 9 perfect guesses. In this case, subjects show more consistency and therefore a better identification of behavioral rules, although we restrict the sample to 71 subjects. However, the results regarding the correlation in panel B are very similar to those in panel A: both non-strategic and *L2* behavioral types show similar likelihood of following a naive (*No-Change*) and a slightly more sophisticated (*Adaptive*) behavioral rule in repeated play.

We therefore conclude that we find no evidence of a correlation between naivete and sophistication in initial play and repeated play.

4 Robustness

One important concern when testing for a correlation between sophistication and naivete in initial and repeated play is that the behavioral type identification is misspecified because some relevant behavioral rules that are relevant to explaining subjects' behavior are not considered. With this concern in mind, we perform two robustness tests. First, we repeat the estimation with elicited behavior in the first part including several alternative behavioral rules in addition to those we already considered. Second, we perform an omitted type specification test to alternatively confirm whether we obtain our result due to the omission of one or many relevant behavioral rules.

4.1 Addition of Alternative Behavioral Rules in Initial Play

We consider 4 alternative behavioral types for the initial play in addition to the 8 we described in Section 2.3. All four types could be considered to be variations of *L1*, where we alter the belief about opponent's behavior. Given that we consider it to be plausible that subjects follow some simple non-strategic rules, it is also plausible that some subjects thought in the same way. Consequently, we consider *L1* as best responding to each of the other non-strategic rules we initially included, that is $L1_{A}$, $L1_{IA}$, $L1_{MaxMax}$ and $L1_{MaxMin}$. Note that these alternative behavioral rules are clearly strategic and closer in spirit to *L2* in terms of strategic sophistication, as they predict a particular opponent's strategy and best responding to that strategy. Additionally, as shown in Table A2 in the Appendix, these additional behavioral types show good separation from the types we initially considered.

As seen in Table A3 in the Appendix, the alternative models appear to show some relevance, although they do not alter the identified type distribution substantially. First, as expected the new alternative behavioral rules steal frequency mostly from *L2*, and the non-strategic types (mostly A). The additional behavioral model that appears to be the most relevant is $L1_{MaxMax}$, which is followed by 9% of subjects. The contingency table displayed in Table 6 shows that subjects following these alternative models are best explained by *No-Change*, followed by *Adaptive* and only a minority is best explained by *Sophisticated* in repeated play. In summary, the consideration of additional alternative behavioral rules to explain initial play does not alter the main results: we find no evidence of correlation between naivete and sophistication in initial and repeated play.

Table 6: Contingency Table with Additional Alternative Behavioral Rules: All Subjects

	Second Part Model				
First Part Model	*No-Change*	*Adaptive*	*Sophisticated*	*Sophisticated 2*	No. of Subjects
Non-strategic	0.53	0.31	0.11	0.06	104
A	0.50	0.30	0.15	0.02	20
IA	0.50	0.25	0.17	0.08	12
$MaxMax$	0.42	0.53	0.00	0.05	19
$MaxMin$	0.65	0.26	0.04	0.04	23
L1	0.53	0.23	0.17	0.07	30
Alternative Models	0.54	0.27	0.12	0.07	41
$L1_A$	0.73	0.18	0.00	0.09	11
$L1_{IA}$	0.57	0.29	0.14	0.00	7
$L1_{MaxMax}$	0.45	0.32	0.14	0.09	22
$L1_{MaxMin}$	0.00	0.00	1.00	0.00	1
L2	0.66	0.24	0.06	0.04	50
L3	-	-	-	-	0
NE	0.67	0.00	0.33	0.00	3
No. of Subjects	112	55	20	11	198

Notes: The table shows the proportion of subjects identified as each of the behavioral rules of the first part that were also classified as each of the behavioral rules of the second part.

4.2 Specification Test: Omitted Types

Similar in spirit to the previous robustness test, we also perform an omitted type specification test (as in Costa-Gomes and Crawford, 2006) to rule the possibility that we did not consider relevant models.

In this test, instead of proposing alternative behavioral models, we let the actual subject behavior in our sample inform us of potential alternative rules. If we left out a rule that actually complies with subjects' behavior, we would expect that some of the subjects behave similarly to this rule. Therefore, we consider the observer behavior as potential new rules in the following manner. In addition to all 12 behavioral rules considered in the previous section we add each subject's actual behavior as an additional behavioral rule, one subject at a time, and re-estimate the mixture-of-types model as many times as the number of subjects in our population, that is, 198 times. While conducting this exercise, we check whether the added subject's behavioral rule is able to explain other subjects' behavior *better* than the existing 12 models and whether the rule can attract sufficient relevance, where we impose a threshold of 15% of the population frequency.

We find three such of those subjects (subject numbers 31, 85, and 86). What strategies are these subjects following? First, we check for similarity of these subjects' behavior (or alternatively, separation). These subjects appear to reflect the same type of behavior as they show very little separation (0.21 between the behavior of subject 31 and subject 85, 0.14 between the behavior of subject 31 and subject 86, and 0.36 between the behavior of 85 and 86). Second, we check their separation from other existing behavioral rules, as shown in Table A4 in the Appendix. All three behavioral rules are well separated from all other considered rules, with the exception of *L2*, showing a separation equal or inferior to 0.43. Third, consistent with this finding, we also observe that when we consider these alternative models in the mixture-of-types model estimation, the behavioral rule that loses the most frequency is indeed *L2*. Finally, we directly consider the actions of these subjects and find that their behavior is mostly consistent with *L2* but in a few games mimic *L1*.[5]. In particular, the strategy profile of subjects 85 and 31 diverge from *L2* or *L1* behavior in only two decisions,

[5] In particular, the strategy profile of subject 31 is: 3 3 3 3 1 1 3 3 3 1 3 2 2 1; the strategy profile of subject 85 is: 3 3 3 1 1 3 3 1 3 1 3 2 3 1; and the strategy profile of subject 86 is: 3 3 3 3 1 1 3 1 3 1 3 2 2 1.

and that of subjects 86 diverges in only three decisions.

We conclude that these subjects show some variation from the existing *L2* behavioral type; however none of them obtains a population frequency higher than that of *L2* when incorporated into the estimation together, as shown in Table A5, or one by one.

Does the result of the correlation between the sophistication and naivete between initial and repeated play change when these new empirically motivated behavioral rules are considered? Table 7 shows that subjects following these alternative models are best explained by *No-Change*, followed by adaptive learners, with very similar proportions as those in Table 5. Therefore, we again conclude that we find no evidence for any correlation between naivete and sophistication in initial and repeated play.

Table 7: Contingency Table with the Addition of Three Subjects' Behavioral Rules: All Subjects

	Second Part Model				
First Part Model	*No-Change*	*Adaptive*	*Sophisticated*	*Sophisticated 2*	No. of Subjects
Non-strategic	0.54	0.32	0.09	0.04	74
A	0.73	0.18	0.09	0.00	11
IA	0.33	0.33	0.17	0.17	6
$MaxMax$	0.36	0.57	0.00	0.07	14
$MaxMin$	0.65	0.25	0.05	0.05	20
L1	0.52	0.30	0.17	0.00	23
Alternative Models	0.45	0.29	0.16	0.10	31
$L1_A$	0.50	0.33	0.00	0.17	6
$L1_{IA}$	0.50	0.33	0.17	0.00	3
$L1_{MaxMax}$	0.44	0.28	0.17	0.11	18
$L1_{MaxMin}$	0.00	0.00	1.00	0.00	1
Subject 31	0.69	0.06	0.13	0.13	16
Subject 85	0.65	0.25	0.10	0.00	20
Subject 86	0.57	0.29	0.10	0.05	21
L2	0.62	0.29	0.03	0.06	34
L3	-	-	-	-	0
NE	0.50	0.00	0.50	0.00	2
No. of Subjects	112	55	20	11	198

Notes: The table shows the proportion of subjects identified as each of the behavioral rules of the first part that were also classified as each of the behavioral rules of the second part.

5 Discussion

In this paper we have explored the relationship between the sophistication and naivete of models in initial and repeated play. Is a naive player in initial play more likely than a more sophisticated player to use a naive model in repeated play? We use an experimental design and mixture-of-types model econometric estimation to answer this empirically motivated research question.

Consistent with previous findings, we find that the Nash equilibrium is not well suited to explain the initial responses of individuals. The non-equilibrium rules that best explain individual behavior appear to be level-*2*, level-*1* and altruistic type of thinking. Additionally, consistent with previous findings, adaptive behavior appears to be quite common in repeated play, although the majority simply ignores the feedback and repeats the previously used strategy. Addressing the central question, exploiting the within-subject design, we find that the behavior in repeated play is independent of the behavior in initial play, so we conclude that naivete and sophistication in initial play are not related to naivete and sophistication in repeated play.

The main result of our paper is reminiscent of the results of Costa-Gomes and Weizscker (2008) and Knoepfle et al. (2009). The former found an inconsistency between the behavior shown by actions and elicited beliefs regarding opponents' expected behavior. The latter found that eye-tracking results favor much more sophisticated learning than do actual decision data, again finding an inconsistency between the two. It could indeed be the case that, similar to actions and beliefs or actions and eye-tracking, individuals treat initial and repeated play as different and independent tasks.

References

[1] Brañas-Garza, P., Espinosa, M. P., and P. Rey-Biel (2011). "Travelers types". *Journal of Economic Behavior and Organization*, 78(1-2), 25-36.

[2] Camerer, C., and T.-H. Ho (1999). "Experience-weighted attraction learning in normal form games." *Econometrica*, 67(4), 827-874.

[3] Camerer, Colin F., T.-H. Ho, and J.-K. Chong (2002). "Sophisticated experience-weighted attraction learning and strategic teaching in repeated games." *Journal of Economic theory* 104.1: 137-188.

[4] Camerer, C. F., Ho, T. H., and J. K. Chong (2004). "A cognitive hierarchy model of games". *The Quarterly Journal of Economics*, 119(3), 861-898.

[5] Cheung, Y. W., and Friedman, D. (1997). "Individual learning in normal form games: Some laboratory results". Games and economic behavior, 19(1), 46-76.

[6] Conlisk, J. (1993a). "Adaptation in games: Two solutions to the Crawford puzzle". Journal of Economic Behavior and Organization, 22(1), 25-50.

[7] Conlisk, J. (1993b). "Adaptive tactics in games: Further solutions to the Crawford puzzle". Journal of Economic Behavior and Organization, 22(1), 51-68.

[8] Costa-Gomes, M. A., V. P. Crawford, and B. Broseta (2001). "Cognition and behavior in normal-form games: An experimental study," *Econometrica*, 69: 1193-1235.

[9] Costa-Gomes, M. A., and V. P. Crawford (2006). "Cognition and behavior in two-person guessing games: An experimental study." *The American Economic Review*, 96.5: 1737-1768.

[10] Costa-Gomes, M. A., and G. Weizscker (2008). "Stated beliefs and play in normal-form games". *The Review of Economic Studies*, 75(3), 729-762.

[11] Crawford, V.P., M. A. Costa-Gomes and N. Iriberri (2013). "Structural Models of Non-equilibrium Strategic Thinking:Theory, Evidence, and Applications," *Journal of Economic Literature*, 51:1: 5-62.

[12] Eckel, Catherine C. and P. J. Grossman (2002). "Sex differences and statistical stereotyping in attitudes toward financial risk." *Evolution and Human Behavior*, 23(4), 281295.

[13] Erev, I., and A. E. Roth (1998). "Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria". *The American Economic Review*, 848-881.

[14] Feltovich, N. (2000). "Reinforcementbased vs. Beliefbased Learning Models in Experimental Asymmetricinformation Games". *Econometrica*, 68(3), 605-641.

[15] Fischbacher, U. (2007). "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics*, 10(2), 171-178.

[16] Frederick, S. (2005). "Cognitive reflection and decision making." *Journal of Economic perspectives*, 19(4), 25-42.

[17] Fudenberg, D., and D. Levine. (1998a). "Learning in games." *European Economic Review* 42.3-5: 631-639.

[18] Fudenberg, D., and D. Levine. (1998b). Theory of learning in games. Cambridge, MA: MIT Press.

[19] García-Pola, B., Iriberri, N. and J. Kovářík (2016). "Non-Equilibrium Play in Centipede Games". CEPR Discussion Papers 11477, C.E.P.R. Discussion Papers.

[20] Goeree, J. K., and C. A. Holt (2001). "Ten little treasures of game theory and ten intuitive contradictions." *The American Economic Review* 91.5: 1402-1422.

[21] Greiner, B. (2015). "Subject pool recruitment procedures: organizing experiments with ORSEE". *Journal of the Economic Science Association*, 1(1), 114-125.

[22] Ho, T.H., Park, S.E., and X. Su (2015). "A Bayesian Level-k Model in n-Person Games." Mimeo.

[23] Iriberri, N., and P. Rey-Biel (2013). "Elicited beliefs and social information in modified dictator games: what do dictators believe other dictators do". *Quantitative Economics* 4: 515-547.

[24] Knoepfle, D. T., Wang, J. T. Y., and C. F. Camerer (2009). "Studying learning in games using eye-tracking". *Journal of the European Economic Association*, 7(2-3), 388-398.

[25] McKelvey, R. D., and T. R. Palfrey (1995). "Quantal Response Equilibrium for Normal-Form Games". *Games and Economic Behavior* 10: 6–38.

[26] Milgrom, P., and J. Roberts (1991). "Adaptive and sophisticated learning in normal form games". *Games and economic Behavior*, 3(1), 82-100.

[27] Nagel, R. (1995). "Unraveling in guessing games: An experimental study." *The American Economic Review* 85.5 : 1313-1326.

[28] Rey-Biel, P. (2009). "Equilibrium play and best response to (stated) beliefs in normal form games". *Games and Economic Behavior*, 65(2), 572-585.

[29] Selten, R. (1991). "Anticipatory learning in two-person games." In Game equilibrium models I (pp. 98-154). Springer, Berlin, Heidelberg.

[30] Sobel, J. (2000). "Economists' models of learning." *Journal of Economic Theory* 94.2: 241-261.

[31] Stahl, D. O., and Wilson, P.W. (1994). "Experimental evidence on players' models of other players." *Journal of Economic Behavior and Organization* 25.3: 309-327.

[32] Stahl, D. O., and Wilson, P.W. (1995). "On players' models of other players: Theory and experimental evidence." *Games and Economic Behavior* 10.1: 218-254.

[33] Stahl, D. O., (2003). "Sophisticated Learning and Learning Sophistication." Mimeo.

[34] Toplak, M. E., West, R. F., and K. E. Stanovich (2014). "Assessing miserly information processing: An expansion of the Cognitive Reflection Test." *Thinking and Reasoning*, 20(2), 147-168.

A Additional Tables And Figures

Table A1: Summary of Socio-Demographic Variables of the Subject Population

Variables	Mean Values	Stand. Dev.
Men	0.41	
Age	21.73	2.99
Spanish	0.87	
University Entry Grade (out of 10)	6.85	1.16
Distribution over Field of Study:		
Social Science	0.77	
Applied Science	0.17	
Natural Science	0.04	
Distribution over risk choices:		
1.5€with 0.50 or 1.5€with 0.50	0.31	
1.3€with 0.50 or 1.8€with 0.50	0.11	
1.1€with 0.50 or 2.1€with 0.50	0.26	
0.9€with 0.50 or 2.4€with 0.50	0.07	
0.7€with 0.50 or 2.7€with 0.50	0.04	
0.6€with 0.50 or 2.8€with 0.50	0.04	
0.4€with 0.50 or 2.9€with 0.50	0.02	
0€with 0.50 or 3€with 0.50	0.16	
Cognitive reflection test:		
Percent of correct in cognitive reflection test: Q1	0.28	
Percent of correct in cognitive reflection test: Q2	0.17	
Percent of correct in cognitive reflection test: Q3	0.41	

Notes: *Men* takes the value of 1 if the subject is male. *Age* reflects the age in years. *Spanish* takes the value of 1 if the subject is Spanish. *University Entry Grade* is normalized to a grade out of 10. *Social Science*, *Applied Science* and *Natural Science* take the value of 1 if the subject is studying a social, applied or natural science. Risk Choice was elicited via Eckel and Grossman (2002), where choices are ordered from the safest to riskiest. Finally, the cognitive reflection test includes questions from Toplak et al. (2014) designed to avoid the possibility that the original test from Frederick (2005) is already known by the subjects. The questions are the following: 1. If John can drink one barrel of water in 6 days, and Mary can drink one barrel of water in 12 days, how long would it take them to drink one barrel of water together? (correct answer 4 days; intuitive answer 9); 2. Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are in the class? (correct answer 29 students; intuitive answer 30); 3. A man buys a pig for \$60, sells it for \$70, buys it back for \$80, and sells it finally for \$90. How much has he made? (correct answer \$20; intuitive answer \$10).

Table A2: Separation of Different Behavioral Rules with Additional Alternative Behavioral Models

	A	IA	$MaxMax$	$MaxMin$	$L1$	$L1_A$	$L1_{IA}$	$L1_{MaxMax}$	$L1_{MaxMin}$	$L2$	$L3$	NE
A	0.00											
IA	0.60	0.00										
$MaxMax$	0.46	0.62	0.00									
$MaxMin$	0.71	0.65	0.57	0.00								
$L1$	0.57	0.76	0.50	0.79	0.00							
$L1_A$	0.25	0.62	0.36	0.64	0.64	0.00						
$L1_{IA}$	0.64	0.98	0.64	0.71	0.50	0.57	0.00					
$L1_{MaxMax}$	0.75	0.73	0.64	0.57	0.79	0.50	0.64	0.00				
$L1_{MaxMin}$	0.71	0.65	0.93	0.64	0.71	0.71	0.64	0.57	0.00			
$L2$	0.75	0.58	0.57	0.64	0.71	0.57	0.79	0.5	0.79	0.00		
$L3$	0.86	0.80	0.86	0.64	0.79	0.71	0.43	0.50	0.50	0.71	0.00	
NE	0.57	0.51	0.86	0.64	0.79	0.57	0.50	0.86	0.57	0.79	0.57	0.00

Notes: The table reports the proportion of strategies across all 14 games in which the different behavioral models predict different strategies. The minimum possible separation value is 0, when the two models prescribe the same strategy in all 14 games, and the maximum possible separation value is 1, when two models predict a different strategy in each of the 14 games.

Table A3: Behavioral Type Identification for Initial Play: Additional Behavioral Types

	Minimum Number of Perfect Guesses							
	No constraints		7		9		11	
	Main	Alt.	Main	Alt.	Main	Alt.	Main	Alt.
A	0.15	0.10	0.13	0.10	0.10	0.07	0.14	0.11
IA	0.07	0.06	0.06	0.05	0.01	0.01	0.00	0.00
$MaxMax$	0.11	0.10	0.11	0.10	0.09	0.07	0.05	0.04
$MaxMin$	0.12	0.12	0.11	0.11	0.11	0.09	0.10	0.07
$L1$	0.16	0.15	0.17	0.16	0.22	0.18	0.33	0.25
$L1_A$		0.06		0.06		0.04		0.00
$L1_{IA}$		0.04		0.04		0.03		0.04
$L1_{MaxMax}$		0.11		0.11		0.15		0.21
$L1_{MaxMin}$		0.01		0.01		0.01		0.00
$L2$	0.36	0.25	0.37	0.26	0.47	0.35	0.38	0.29
$L3$	0.02	0.00	0.02	0.00	0.00	0.00	0.00	0.00
NE	0.02	0.02	0.02	0.02	0.01	0.01	0.00	0.00
No. of Subjects	198	198	186	189	93	113	21	28

Notes: The table displays the population frequencies estimated for the main specification, shown in Table 3 and when adding alternative models in initial play.

Table A4: Separation of the Three Relevant Subjects' Behavior from other Behavioral Models

	A	IA	$MaxMax$	$MaxMin$	$L1$	$L1_A$	$L1_{IA}$	$L1_{MaxMax}$	$L1_{MaxMin}$	$L2$	$L3$	NE
Subject 31	0.57	0.58	0.71	0.71	0.50	0.57	0.57	0.57	0.64	0.36	0.71	0.57
Subject 85	0.54	0.65	0.50	0.71	0.57	0.50	0.50	0.64	0.79	0.36	0.79	0.64
Subject 86	0.57	0.55	0.64	0.71	0.50	0.57	0.57	0.50	0.64	0.43	0.79	0.71

Notes: The table reports the proportion of strategies across all 14 games in which the three subjects' behavioral models predict different strategies from the rest of the considered models. The minimum possible separation value is 0, when the two models prescribe the same strategy in all 14 games, and the maximum possible separation value is 1, when two models predict a different strategy in each of the 14 games.

Table A5: Behavioral Type Identification for Initial Play: Additional Behavioral Types

	Minimum Number of Perfect Guesses							
	No constraints		7		9		11	
	Main	Alt.	Main	Alt.	Main	Alt.	Main	Alt.
A	0.10	0.06	0.10	0.06	0.07	0.06	0.11	0.07
IA	0.06	0.03	0.05	0.03	0.01	0.01	0.00	0.00
$MaxMax$	0.10	0.07	0.10	0.07	0.07	0.05	0.04	0.02
$MaxMin$	0.12	0.10	0.11	0.09	0.09	0.07	0.07	0.04
L1	0.15	0.12	0.16	0.12	0.18	0.13	0.25	0.16
$L1_A$	0.06	0.03	0.06	0.03	0.04	0.01	0.00	0.00
$L1_{IA}$	0.04	0.03	0.04	0.03	0.03	0.02	0.04	0.02
$L1_{MaxMax}$	0.11	0.09	0.11	0.09	0.15	0.11	0.21	0.13
$L1_{MaxMin}$	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00
L2	0.25	0.17	0.26	0.017	0.35	0.21	0.29	0.18
L3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
NE	0.02	0.01	0.02	0.01	0.01	0.01	0.00	0.00
Subject 31		0.08		0.08		0.09		0.07
Subject 85		0.10		0.10		0.11		0.16
Subject 86		0.11		0.11		0.12		0.16
No. of Subjects	198	198	189	195	113	142	28	45

Notes: The table displays the population frequencies estimated for the main specification, shown in Table 3 and when adding alternative models in initial play.

B Mixture-of-types Likelihood Function

We assume that a subject i employing rule k makes type-k's decision with probability $(1-\varepsilon_i)$, but makes a mistake with probability $\varepsilon_i \in [0,1]$. In such a case, she plays each of the three available strategies uniformly at random. As in most mixture-of-types model applications, we assume that the errors are identically and independently distributed across games and subjects and that the errors are subject-specific (as in for example Iriberri and Rey-Biel, 2013). The first assumption facilitates the statistical treatment of the data, while the second considers that some subjects may be noisier and thus make more error than others.

The likelihood of a particular individual of a particular type can be constructed as follows. Let $P_k^{g,j}$ be type-k's predicted choice probability for strategy j in game g. Some rules may predict more than one strategy in a particular game. This characteristic is reflected in the vector $P_k^g = (P_k^{g,1}, P_k^{g,2}, P_k^{g,3})$ with $\sum_j P_k^{g,j} = 1$.

For each individual in each game, we observe the choice and whether it is consistent with k. Let $x_i^{g,j} = 1$ if strategy j is chosen by subject i in game g in the experiment and $x_i^{g,j} = 0$ otherwise. The likelihood of observing a sample $x_i = (x_i^{g,j})_{g,j}$ given type k and subject i is then

$$L_i^k(\varepsilon_i|x_i) = \prod_g \prod_j \left[(1-\varepsilon_i)P_k^{g,j} + \frac{\varepsilon_i}{3}\right]^{x_i^{g,j}}. \tag{1}$$

Finally, the likelihood function is given by the sum of all behavioral types that are considered.

$$L_i(\varepsilon_i|x_i) = \sum_k p_i L_i^k(\varepsilon_i|x_i) \tag{2}$$

p_i takes a value of 1 for the behavioral type k that best explains the individual behavior and 0 for the rest of the considered behavioral types.

For explaining initial play, we consider $K = 8$ behavioral types or models: *A*, *IA*, *MaxMax*, *MaxMin*, *L1*, *L2*, *L3* and *NE*. To explain repeated play with feedback, we consider $K = 4$ different behavioral types: *No-Change*, *Adaptive*, *Sophisticated* and *Sophisticated 2*.

C Translation of Instructions

The original instructions were in Spanish. We provide a translation of instructions into English

THANK YOU FOR PARTICIPATING IN OUR EXPERIMENT!

We will now start the experiment. From now on, you are not allowed to speak, look at what other participants do or walk around the room. Please turn off your phone. If you have any questions or need help, raise your hand and one of the researchers will talk with you. Please, do not write on these instructions. If you do not follow these rules, YOU WILL BE ASKED TO LEAVE THE EXPERIMENT AND NO PAYMENT WILL BE GIVEN TO YOU. Thank you.

The university and the research projects have provided the funds for the realization of this experiment. You will receive 3 Euros for having arrived on time. Additionally, if you follow the instructions correctly you have the possibility to earn more money. This is a group experiment. The amount you can earn depends on your decisions, the decisions of other participants, and chance. Different participants can earn different amounts.

No participant will be able to identify another by their decisions or by their profits in the experiment. The researchers will be able to observe the profits of each participant at the end of the experiment, but we will not associate the decisions you have made with the identity of any participant.

EARNINGS:

During the experiment you can earn experimental points. At the end, each experimental point will be exchanged for Euros, and exactly 1 experimental point is worth 0.5 Euros. Everything you win will be paid in cash in a strictly private way at the end of the experimental session.

Your final earnings will be the sum of the 3 Euros you receive for participating plus what you earn during the experiment.

Each experimental point equals 50 cents, so 2 experimental points equals 1 Euro ($2x0.5 = 1$ Euro).

If, for example, you earn a total of 20 experimental points you will receive a total of 13 Euros (3 Euros as payment for participation and 10 Euros from the conversion of the 20 experimental points to Euros).

If, for example, you earn 4 experimental points you will obtain 5 Euros ($4x0.5 = 2$ and $2 + 3 = 5$).

If, for example, you earn 44 experimental points you will obtain 25 Euros ($44x0.5 = 22$ and $22 + 3 = 25$).

PARTS OF THE EXPERIMENT:

The experiment consists of two parts. You will participate by operating a computer. In the first part there will be 14 rounds, where you will make 14 decisions. In the second part, there will also be 14 rounds, where you will make 14 decisions. At the end of the experiment, when you have completed the two parts of the experiment, the computer will randomly choose two of the 28 rounds, and you will be paid for the money you received in those two rounds chosen at random, plus the 3 Euros for participating.

Before beginning each part of the experiment, we will explain in detail what kind of decisions you can make and how you can obtain experimental points.

When we are all ready, we will start the first part of the experiment by explaining the instructions of the first part of the experiment in detail.

FIRST PART OF THE EXPERIMENT:

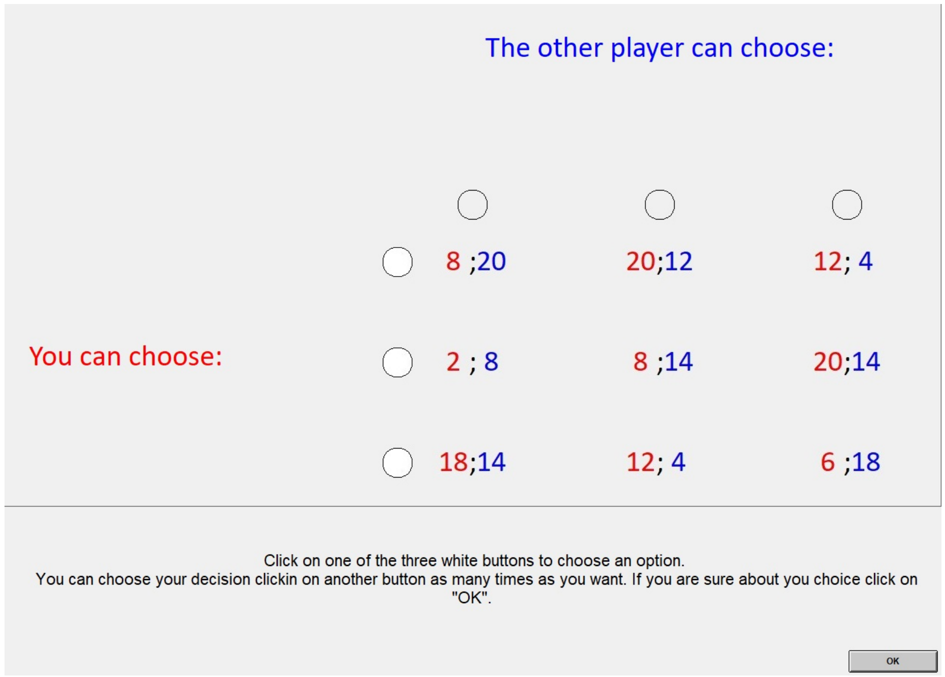
The first part of the experiment consists of 14 rounds. In each of the 14 rounds, you will be paired with a participant chosen at random from this session. The other participant will be different in each of the rounds, so you will never be paired with the same participant more than once. From now on, we will refer to you as "You" and the other participant as "other participant".

In each round you will have to make a decision, choosing among three possible options. Each decision will be presented in the form of a table similar to the one below (but with different values). You will see the corresponding table each time you have to choose an option. Each row of the table corresponds to an option that you can choose. The decision you must make is to choose one option. The other participant will also have to choose, independently of you, among their options, which correspond to the columns of the table. That is, you choose among rows, while the other participant chooses among columns. However, to simplify things, the experiment is programmed in such a way that all the participants - including the person with whom you are matched - see their decision as shown in the example. That is, each of you will be presented with your possible actions in the rows of the table, and your experimental points will be shown in red. At the time of choosing, you will not know the option chosen by the other participant, and when the other participant is choosing their option, they will not know the option that you have chosen.

The number of experimental points you earn in each of the rounds depends on the option you have chosen and the option that the other participant has chosen.

The table of experimental points you see below is an example of what you will see in each of the rounds.

Example:



For example, if this round is chosen at random and you select the first option (row) and the other participant select the second option (column), you will obtain 20 experimental points and the other participant will receive 12 experimental points.

Another example: if this round is chosen at random and you select the third option (row) and the other participant selects the first option (column), you will obtain 18 experimental points and the other participant will receive 14 experimental points.

These are just two examples to better understand how decisions affect the experimental points you can earn and do not intend to suggest what decisions you should make.

To make a selection, click on the white button next to the desired. Then, the button will turn red to indicate which option you have selected. Once you have chosen an option, the choice is not final and you can change your selection as many times as you want by clicking on another button, until you press the “OK” button that will appear in the lower right corner of each screen. Once you have clicked “OK” the selection will be final and you will proceed to the next round. You will not be able to move to the next round until you have chosen an option and clicked “OK”. You will not have any time restrictions. Take as much time as you need in each round. When all of you have made your decisions in each of the 14 rounds, we will explain the second part of the experiment.

Summary:

- Your experimental points will be shown in red, and the experimental points of the other participant will be shown in blue.
- You will participate in 14 different rounds. In each of the rounds the table of experimental points will be different and you will be paired with a different participant chosen at random from this session.
- In each round, you can choose among three different options (rows) and the experimental points that you earn depend on the option you select, the option that the other participant selects, and whether that round is chosen at random at the end of the experiment.

We will start the first part of the experiment in a few moments. Before starting the first part, you will see a new example and you will have to answer several questions. If you have any questions or need help at any time during the experiment, please raise your hand and one of the investigators will talk to you.

SECOND PART OF THE EXPERIMENT:

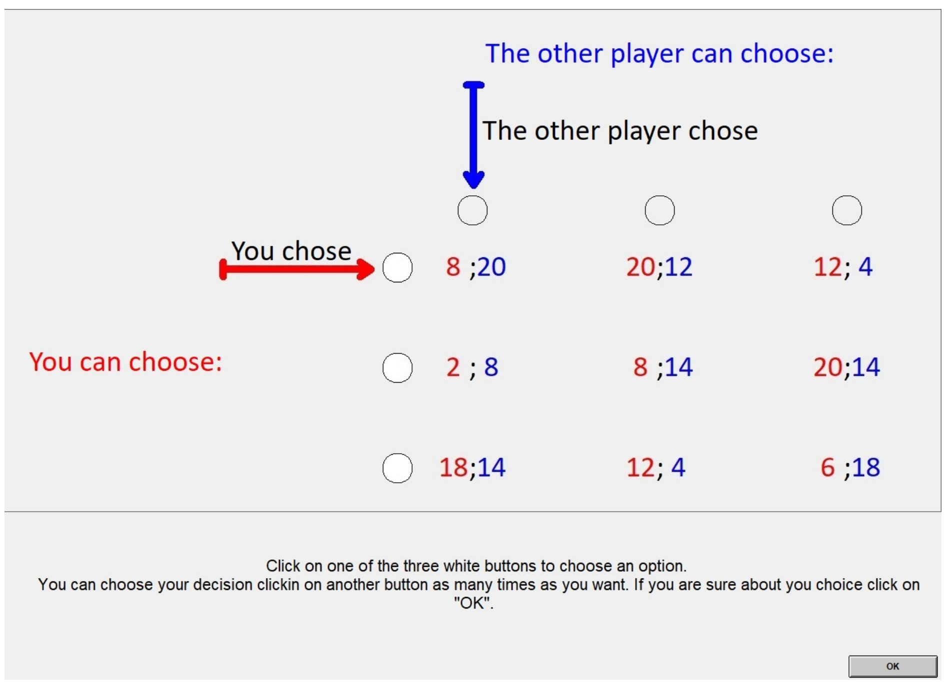
The second part of the experiment also consists of 14 rounds and will work in a similar way to the first part. That is, the tables of experimental points that you will see in each of the 14 rounds in this second part will be the same as those you saw in the first part of the experiment. As in the first part, in each of the 14 rounds, you will be paired with a participant chosen at random from this session. However, in each of the rounds, the other participant with whom you have been paired in this part does not have to be the same as the participant with whom you were paired in the first part. The pairing is performed again at random. In each of the rounds, the other participant, chosen at random, will be different, so you will never be paired with the same participant more than once.

As in the first part, both you and the other participant can choose among three possible options. The experimental points that you can earn in each of the rounds depend on the option that you select and the option that the other participant selects, as well as on whether that particular round is chosen at random at the end of the experiment.

Unlike the first part, when you see the table of experimental points, you can also observe the option that you chose in the first part and the option that was chosen in the first part by the participant with whom you are paired in this part. The option that you both chose in the first part will be indicated by an arrow and will say "You chose" and "The other chose". The information you observe will be the same for the participants with whom you are paired.

The table of experimental points you see is an example of what you will see in each of the rounds.

Example:



As in the first part, for example, if this round is chosen at random and you select the first option (row) and the other participant selects the second option (column), you will earn 20 experimental points and the other participant will earn 12 experimental points.

Another example: if this round is chosen at random and you select the third option (row) and the other participant selects the first option (column), you will obtain 18 experimental points and the other participant will receive 14 experimental points.

These are just two examples to better understand how decisions affect the experimental points you earn and are not intended to suggest what decisions you should make.

Unlike the first part, in this part of the experiment, you can observe, as indicated in the example, which option you chose and which option the other participant chose in the first part. For example, in the example table, you chose the second option (row) and the other participant chose the second option (column). The other participant can also observe the option you chose and the option he/she chose: you both have the same information. Now you will have to make a choice again.

You can make your decision in the same way as in the first part, by clicking on the button of the option you want to choose and confirming by pressing "OK". You will not have any time restrictions. Take as much time as you need in each of the rounds. When all of you have made your decisions in each of the 14 rounds, the experiment will end.

Summary:

- Your experimental points will be shown in red and the experimental points of the other participant will be shown in blue.
- You will participate in 14 different rounds. In each round, the table of experimental points will be different and you will be paired with a different participant chosen at random from this session.
- Unlike the first part, you can now see which option you chose in the first part, and which option the other participant chose in the first part. The other participant will also be able to observe the option that he/she chose, as well as the option that you chose.
- In each round, you can choose among three different options (rows) and the experimental points depend on the option you have chosen, on the option chosen by the other participant, and whether that round is chosen at random at the end of the experiment.

We will start the second part of the experiment in a few moments. If you have any questions or need help at any time during the experiment, please raise your hand and one of the investigators will talk to you.