Laboratory of Economics and Management

Sant'Anna School of Advanced Studies
Piazza Martiri della Libertà, 33 - 56127 PISA (Italy)
Tel. +39-050-883-343 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://www.lem.sssup.it/

LEM

Working Paper Series

Feature-based Choice and Similarity in Normal-form Games: An Experimental Study

Giovanna Devetag°
Sibilla Di Guida[§]

° University of Perugia, Italy
[§] CIFREM- Univertity of Trento, Italy

2010/18 February 2013

ISSN (online) 2284-0400

Feature-Based Choice and Similarity Perception in Normal-form Games: An Experimental Study

Sibilla Di Guida[1]
CIFREM, University of Trento, via Inama 5, 38122 Trento, Italy
sibilla.diguida@gmail.com

Giovanna Devetag (corresponding author)
Dipartimento di Discipline Giuridiche e Aziendali, University of Perugia
Via Pascoli 20, 06123 Perugia, Italy, tel/fax n. +39-075-5852174, devetag@unipg.it

June 2012

Abstract

In this paper, we test the effect of descriptive “features” on initial strategic behavior in normal form games, where the term “descriptive” indicates all those features which can be modified without altering the (Nash) equilibrium structure of a game. Our experimental subjects behaved according to some simple heuristics based on descriptive features, and we observed that these heuristics were stable even across strategically different games. These findings indicate the need to incorporate descriptive features into models describing strategic sophistication in normal form games. Analysis of choice patterns and individual behavior indicates that non-equilibrium choices may derive from incorrect and simplified mental representations of the game structure, rather than from beliefs in other players' irrationality. We suggest how level-k and cognitive hierarchy models might be extended to account for heuristic-based and feature-based behavior.

Keywords: normal form games, one-shot games, response times, dominance, similarity, categorization, focal points, individual behavior
JEL classification codes: C72, C91, C92

---

[1] Present address: ECARES, Universitè Libre de Bruxelles, Av. Roosevelt 50, CP 114/4, 1050 Brussels, Belgium

1. Introduction

How do people react to a strategic situation they have never encountered before, and which they have to face only once with an unknown partner? Behavioral game theory in the last decade has produced abundant experimental evidence in response to this question, which in turn has informed empirically grounded modeling efforts. The enterprise has led to the development of new equilibrium concepts (e.g., QRE, Cursed Equilibrium, to name a few) and of models that take on board individuals' limited reasoning abilities (i.e., the vast family of the "level-k" models: Stahl and Wilson 1994, 1995; Nagel 1995; Ho et al. 1998; Costa-Gomes et al. 2001; Bosch-Domènech et al. 2002; Crawford 2003; Camerer et al. 2004; Costa-Gomes and Crawford 2006; Crawford and Iriberri, 2007).

Behavioral models of game playing, although relaxing the most implausible assumptions of standard game theory, still rely on the notion of strategic thinking, i.e., they assume that players form beliefs about other players (in other words, players have a mental model of the other player's strategic behavior) and try to maximize their utility given these beliefs.

Some of the available evidence, however, suggests a partly different picture of behavior in one-shot games, by pointing out systematic and widespread inconsistencies between choices and beliefs (Costa-Gomes and Weizsäcker 2008), by suggesting, more radically, that players may use decision heuristics that do not imply any beliefs at all (Weizsäcker 2003), and finally, by hinting at the possibility that players may even act on the basis of a very simplified and/or incorrect model of the "true" strategic situation (Devetag and Warglien 2008, Rydval et al., 2009).

While it is certainly beyond the scope of this paper to provide an univocal answer to the question of how people behave in one-shot interactions, we submit that players in several cases may use choice heuristics that are based on a simplification of the original decision problem, obtained either by ignoring the other players' motivations or by ignoring a subset of the game outcomes. These heuristics have the property of being "vulnerable" to the influence of features other than a game inner strategic structure: as a consequence, predictable changes in observed behavior can occur as a function of the presence vs. absence of these equilibrium-irrelevant features.

More specifically, we argue that players in reasonably complex single-shot games, with no opportunity for learning and no feedback, at first look for "obvious" and "natural" solutions to the strategic problem they face: one such natural solution is picking a strategy which is both attractive and safe, i.e., one with high payoff sum and low payoff variance. Alternatively, an equally "natural" solution is selecting the strategy supporting an outcome that is attractive for both players, which we call focal point. The first behavior is compatible with "level 1" types in level-k models of strategic thinking (Camerer et al., 2004; Costa-Gomes et al., 2001; Stahl and Wilson, 1995) and may derive either from diffuse priors on the opponent's play or from a tendency to ignore opponents' moves entirely (Costa-Gomes and Weizsäcker, 2008; Weizsäcker, 2003). However, unlike the above-mentioned models, we assume that payoff variance (taken as an intuitive measure of the risk involved in choosing a strategy) plays an important role in determining "level 1" behaviors. To the best of our knowledge, we are the first to test the role of payoff variance in influencing behavior in games. The second solution is strategic because it relies on Schelling salience (Mehta et al., 1994; Sugden, 1993) which has been identified in experiments on matching games and which is known to be very effective in promoting coordination. However, we call this strategic approach "naive", in that focal points in our games are not equilibrium outcomes.

Only in the absence of features suggesting attractive "solutions" may players start to reason more strategically in a game-theoretic sense, and in doing so find their way to equilibrium.

We also hypothesize that games sharing features such as the presence of a safe-and-attractive strategy and a focal point may trigger similar behaviors (at both aggregate and individual levels), despite very different inner strategic structures; conversely, games which differ feature-wise but which present the same equilibrium properties may trigger very different behaviors. Theories of cross-game similarity are crucial when modeling important phenomena such as cross-game transfer and generalization. It is widely acknowledged that the games we play in real life are at most similar to each other but never identical (unlike the typical "Groundhog day" lab situation) and, as many of our decision processes are case-based or analogy-based (Gilboa and Schmeilder, 1995; Jehiel, 2005), it becomes essential to understand when players perceive two games as being similar. Surprisingly, there are very few studies investigating cross-game similarity perception. Among these, Knez and Camerer (2000) test transfer of precedent between a Prisoner's Dilemma and a Weak Link game, and introduce the distinction between surface (or descriptive) and structural similarity. In their design, transfer of precedent is triggered only in the presence of descriptively similar features between the two games (such as action labeling). Rankin et al. (2000) tested coordination behavior in perturbed environments by having subjects play a series of stag hunt games with randomly perturbed payoffs and action labels, and found that, when descriptive similarity is impeded, convergence to the payoff-dominant equilibrium is more frequent. Hence, understanding what features are relevant in eliciting similarity perception between games is crucial for modeling both repeated behaviors in ever-changing environments and phenomena of generalization from experience.

Showing that choice behavior responds to theoretically irrelevant features is obviously not a novel approach, and early examples date back to the famous studies of “framing” effects in individual decision making (Kahneman and Tversky 1979; Tversky and Kahneman 1981). However, since our manipulations are “economic” in nature, implying exclusively changes in payoffs and, for one game only, changes in the position of the focal point in the matrix, they resemble more studies like Goeree and Holt (2001). Our feature manipulations influence behavior significantly and predictably, suggesting that players respond to features in the way hypothesized. We also show that players respond similarly to games that are similar in terms of features, even when they belong to very different strategic types. Hence, a classification of games based on descriptive features (e.g., an outcome with symmetric, high payoffs, a strategy with high expected value and low variance, etc.) turns out to be more useful in predicting initial behavior than a categorization based on a game equilibrium structure.

Our findings are connected with previous studies in several ways: first, they provide evidence of behavior in single-shot normal form games which cannot be adequately explained by equilibrium concepts or by behavioral models which assume a distribution of game-invariant player types and thinking steps. In doing this, they point to the role of strategy variance as an important determinant of choice. Second, they extend the notion of a “focal point” well beyond equilibrium outcomes in symmetric games, showing that focality may be a much more general property of game outcomes, both symmetric and asymmetric.

More generally, our results show that mild payoff changes induce relevant modifications in behavior, which can be parsimoniously explained by the use of decision heuristics that are based on incomplete information processing, in line with suggestions from previous experiments. We argue that these findings may constitute the basis for a theory of similarity that takes into account both structural and descriptive dimensions to describe players’ cross-game similarity perceptions, the latter being more prominent than the former when initial or single-shot behavior is concerned.

Finally, our results add insights to the so-called “pre-game theory” (Camerer, 2003), i.e., they contribute to our understanding of strategic interaction situations as these are perceived and interpreted by the decision makers.

2. The games

In order to test our hypotheses, we used 30 3x3 games in normal form belonging to five well-known game types. The payoff matrices used in the experiment are listed in Table 1.

We selected 5 game types and created 6 versions of each game. In some cases, new Nash equilibria emerged together with the original ones, which always remained.
The chosen basic games were: a game with a strictly dominant strategy for the column player (henceforth, DomCol game); a game without pure strategy Nash Equilibria (noNE), a game with a single pure strategy Nash Equilibrium but not solvable through iterated elimination of dominated strategies (UniqNE), a 3x3 game with weakly dominant strategies similar to a Prisoners' Dilemma (PD)[2], and a Weak Link coordination game (WL).
Our main goal was to examine how the presence or absence of focal points affected subjects' perception of cross-game similarity and strategic behavior, as well as the effect of increasing the variance of the strategy with the highest average payoff (henceforth HA; three levels of variance were introduced: low, medium, high). For this purpose, and in order to identify both their separate and joint effects, we created a matrix for every possible combination of features. Six matrices were created for each basic game: with focal point and HA with low variance; with focal point and HA with medium variance; with focal point and HA with high variance; without focal point and HA with low variance; without focal point and HA with medium variance; without focal point and HA with high variance.

For each game, we identified the strategy with the highest average payoff (HA), the equilibrium strategy (EQ, whenever a pure strategy Nash Equilibrium is present), and a strategy leading to a Focal Point (FP). A Focal Point is any cell containing Pareto-efficient and symmetric payoffs, located at the center of the matrix, except in the Weak Link game, where all symmetric cells were positioned along the main diagonal from the highest to the lowest payoff. Except in a few cases, our Focal Points are not equilibria[3].

---

[2] A referee rightly pointed out that our “modified PD” game is not a Prisoner’s Dilemma, not only because it is a 3x3 game, but also because it does not contain strictly dominant strategies. We decided to name it PD for brevity, simplicity and because the game nonetheless incorporates the basic tension between a cooperative strategy and a (weakly dominant) defecting strategy typical of a PD; however, we agree with the referee that the name is somewhat improper.

[3] Our definition of a focal point differs both from that of Schelling (1963) and from those previously used in all experimental games (Bosch-Domènech and Vriend, 2008; Crawford et al.,

To facilitate the exposition, we called each matrix by the acronym identifying the game type, and by two acronyms identifying its features: FP means a matrix with a focal point, XFP a matrix without focal point, and L, M and H the three levels of variance of the strategy with the highest payoff sum[4].

All the differing versions of the same game were created, changing the cell content as little as possible, and always preserving the game equilibrium structure. In a few cases, these changes added new Nash Equilibria in mixed strategies. In extreme cases, two matrices differed by a single cell. To avoid spurious effects due to the position of the strategy in the matrix, we always kept the position of every strategy fixed in the different versions of the same game, the only exception being the Weak Link game.

Except in one matrix (WL_FP_L), the average payoff of the HA strategy was kept unchanged in the different versions of the same game, and only the payoff distribution was modified so as to change the value of payoff variance.

Matrices without focal point were obtained by breaking the symmetry of payoffs and by reducing payoff magnitude. In the case of the Weak Link game, to keep the game equilibrium structure unaltered, we simply moved the focal cell from the top-left cell to a less "focal" position.

In order to measure the impact of every feature, we kept our three strategies of interest separate whenever possible. For example, in the DomCol game, Row 1 identifies the HA strategy, Row 2 the FP strategy, and Row 3 the EQ strategy. This separation was not possible to obtain in the case of the Prisoner's Dilemma, where EQ and HA coincide by definition: in this case a single row is therefore simultaneously the EQ and the HA strategy. In addition, since keeping the three features of interest separate for both players was impossible, we chose to focus our analysis on the behavior of the row players only. Therefore, unless otherwise specified, all descriptions of strategies and matrices deal with the row player's perspective.

Finally, as we are interested in initial behavior only, we implemented a random rematching scheme with no feedback, to avoid learning and "repeated game effects" as much as possible.

2008; Metha et al., 1994; Sugden, 1995), as we define as "focal" any outcome which is Pareto-efficient and yields identical payoffs to the players.

[4] Also, COS is a strategy yielding a constant payoff (present only in the Weak Link game) and DOM is a weakly dominated strategy (present only in the modified Prisoner's Dilemma). Lastly, QES is a quasi-equilibrium strategy (present in the game without pure strategy Nash Equilibria) in the sense explained in section 4 (see discussion of results).

		HA low var						HA middle var						HA high var					
DomCol	**FP**		C1	C2	C3				C1	C2	C3				C1	C2	C3		
		R1	35,20	35,25	35,30	45%	HA	R1	60,20	20,25	25,30	27%	HA	R1	80,20	10,25	15,30	23%	HA
		R2	5,55	80,80	5,85	38%	FP	R2	5,55	80,80	5,85	42%	FP	R2	5,55	80,80	5,85	43%	FP
		R3	10,20	10,15	40,25*	17%	EQ	R3	10,20	10,15	40,25*	32%	EQ	R3	10,20	10,15	40,25*	33%	EQ
				FP	EQ/HA					FP	EQ/HA					FP	EQ/HA		
	XFP		C1	C2	C3				C1	C2	C3				C1	C2	C3		
		R1	35,20	35,25	35,30	80%	HA	R1	60,20	20,25	25,30	48%	HA	R1	80,20	10,25	15,30	33%	HA
		R2	5,55	50,25	5,85	2%	FP	R2	5,55	50,25	5,85	7%	FP	R2	5,55	50,25	5,85	5%	FP
		R3	10,20	10,15	40,25*	18%	EQ	R3	10,20	10,15	40,25*	45%	EQ	R3	10,20	10,15	40,25*	62%	EQ
				XFP	EQ/HA					XFP	EQ/HA					XFP	EQ/HA		
noNe	**FP**		C1	C2	C3				C1	C2	C3				C1	C2	C3		
		R1	35,15	35,20	35,30	52%	HA	R1	55,15	25,20	25,30	37%	HA	R1	75,15	15,20	15,30	20%	HA
		R2	5,45	75,75	10,80	32%	FP	R2	5,45	75,75	10,80	50%	FP	R2	5,45	75,75	10,80	58%	FP
		R3	15,35	5,25	40,20	17%	QES	R3	15,35	5,25	40,20	13%	QES	R3	15,35	5,25	40,20	22%	QES
				FP	QES/HA					FP	QES/HA					FP	QES/HA		
	XFP		C1	C2	C3				C1	C2	C3				C1	C2	C3		
		R1	35,15	35,20	35,30	73%	HA	R1	55,15	25,20	25,30	53%	HA	R1	75,15	15,20	15,30	53%	HA
		R2	5,45	50,25	10,80	7%	XFP	R2	5,45	50,25	10,80	7%	XFP	R2	5,45	50,25	10,80	0%	XFP
		R3	15,35	5,25	40,20	20%	QES	R3	15,35	5,25	40,20	40%	QES	R3	15,35	5,25	40,20	47%	QES
				XFP	QES/HA					XFP	QES/HA					XFP	QES/HA		
UniqNe	**FP**		C1	C2	C3				C1	C2	C3				C1	C2	C3		
		R1	35,10	35,15	35,10	43%	HA	R1	55,10	25,15	25,10	28%	HA	R1	70,10	20,15	15,10	20%	HA
		R2	10,50	70,70	5,75	47%	FP	R2	10,50	70,70	5,75	45%	FP	R2	10,50	70,70	5,75	43%	FP
		R3	5,10	10,5	40,15*	10%	EQ	R3	5,10	10,5	40,15*	27%	EQ	R3	5,10	10,5	40,15*	37%	EQ
				FP	EQ/HA					FP	EQ/HA					FP	EQ/HA		
	XFP		C1	C2	C3				C1	C2	C3				C1	C2	C3		
		R1	35,10	35,15	35,10	75%	HA	R1	55,10	25,15	25,10	68%	HA	R1	70,10	20,15	15,10	47%	HA
		R2	10,50	50,25	5,75	13%	XFP	R2	10,50	50,25	5,75	3%	XFP	R2	10,50	50,25	5,75	12%	XFP
		R3	5,10	10,5	40,15*	12%	EQ	R3	5,10	10,5	40,15*	28%	EQ	R3	5,10	10,5	40,15*	42%	EQ
				XFP	EQ/HA					XFP	EQ/HA					XFP	EQ/HA		
PD	**FP**		C1	C2	C3				C1	C2	C3				C1	C2	C3		
		R1	35,10	35,5	35,35*	87%	EQ/HA	R1	25,10	60,5	20,20*	80%	EQ/HA	R1	15,10	80,5	10,10*	80%	EQ/HA
		R2	10,35	35,35*	5,35	10%	FP	R2	10,35	35,35	5,60	17%	FP	R2	10,35	35,35	5,80	10%	FP
		R3	15,15	35,10	10,35	3%	DOM	R3	15,15	35,10	10,25	3%	DOM	R3	15,15*	35,10	10,15*	10%	DOM
				FP	EQ/HA					FP	EQ/HA					FP	EQ/HA		
	XFP		C1	C2	C3				C1	C2	C3				C1	C2	C3		
		R1	35,10	35,5	35,35*	92%	EQ/HA	R1	25,10	60,5	20,20*	87%	EQ/HA	R1	15,10	80,5	10,10*	68%	EQ/HA
		R2	10,35	35,25	5,35	5%	XFP	R2	10,35	35,25	5,60	5%	XFP	R2	10,35	35,25	5,80	10%	XFP
		R3	15,15	35,10	10,35	3%	DOM	R3	15,15	35,10	10,25	8%	DOM	R3	15,15*	35,10	10,15*	22%	DOM
				XFP	EQ/HA					XFP	EQ/HA					XFP	EQ/HA		
WL	**FP**		C1	C2	C3				C1	C2	C3				C1	C2	C3		
		R1	60,60*	35,45	5,35	57%	FP	R1	60,60*	35,45	5,35	58%	FP	R1	60,60*	35,45	5,35	10%	FP
		R2	45,35	45,45*	35,35	42%	HA	R2	50,35	50,50*	20,35	33%	HA	R2	60,35	60,60*	5,35	72%	HA
		R3	35,5	35,35	35,35*	2%	COS	R3	35,5	35,35	35,35*	8%	COS	R3	35,5	35,35	35,35*	18%	COS
			FP	HA	COS				FP	HA	COS				FP	HA	COS		
	XFP		C1	C2	C3				C1	C2	C3				C1	C2	C3		
		R1	35,35	45,45*	45,35	48%	HA	R2	20,35	50,50*	50,35	38%	HA	R2	5,35	60,60*	60,35	65%	HA
		R2	5,35	35,45	60,60*	48%	XFP	R2	5,35	35,45	60,60*	50%	XFP	R2	5,35	35,45	60,60*	12%	XFP
		R3	35,35*	35,35	35,5	3%	COS	R3	35,35*	35,35	35,5	12%	COS	R3	35,35*	35,35	35,5	23%	COS
			COS	HA	XFP				COS	HA	XFP				COS	HA	XFP		

Table 1: Summary of all experimentally investigated games, grouped by type of game, level of HA variance, and presence of FP. * : pure strategy Nash Equilibria. Frequency of each row is specified.

3. Experimental design and behavioral predictions

3.1 Experimental design and implementation

The experiment was conducted at the Computable and Experimental Economics Lab (CEEL) of the University of Trento, in 5 different sessions, each having 16 subjects. In each session, 12 subjects were randomly assigned the role of row player and 4 the role of column player, for a total of 60 observations for row players and 20 for column players. Roles were fixed throughout the experiment. Subjects made their choices as row or column players in the 30 matrices, and were re-matched randomly at every round with a player of the opposite role.

All games were presented from the viewpoint of the row player.

No feedback regarding opponents' choice or the obtained payoff was revealed until the end of the experiment.

On entering the lab, subjects were assigned randomly to a pc cubicle and to the role of row or column player. They were given a paper copy of the instructions, which was also read aloud by the experimenter. Control questions were administered before starting the experiment, to ensure that the rules of the experiment had been understood. Particular care was taken to make sure that subjects understood how to read a payoff matrix. In the case of incorrect answers, instructions were repeated (for a translated copy of the instructions and control questions, see appendices B, C, and Figure 1).

FIGURE 1

The experiment was computerized using a Z-Tree based software (Fischbacher, 2007), specifically developed for the purpose. The matrices were presented one at a time in random order, which differed across subjects.

In each round, subjects had to select their preferred strategy by typing the corresponding row number (Figure 1 shows a sample of the software interface).

All players' strategies were recorded and matched randomly, but no feedback was given until the end of the experiment.

At each decision stage, a message appeared after 30 seconds inviting subjects to make their choices. On several occasions subjects used more than 60 seconds to make their decision, showing that the suggestion was not perceived as mandatory[5].

The final payment was determined by the outcomes of 5 matrices picked at random. The exchange rate was announced at the end (and had been made explicit to subjects in the instructions)[6].

After the last matrix had been displayed, one subject selected randomly was asked to verify that tags in a jar each reported the numeric code of one of the matrices played. Subsequently, another randomly selected subject was asked to extract 5 tags from the jar to determine which matrices would be used for payments. Then all subjects were paid according to their own and their assigned partners' choices in the 5 selected matrices.

Before leaving the lab, subjects also took some personality tests and were administered the Holt and Laury lottery choices (Holt and Laury, 2002) with real payments (for a translated copy of the test, see Appendix D). Hence, players' final payments were the sum of their earnings from the 5 matrices selected and their winnings from the lottery. Sessions did not last more than 1.5 hours and average earnings equaled 14 Euros. Minimum earnings were equal to 10 Euros and maximum earnings were equal to 17.50 Euros. Average expected payment was calibrated according to the CEEL Lab guidelines.

3.2 Behavioral predictions

Our choice of features has the goal of revealing the use of decision heuristics leading to intuitive solutions. We aim to extend our implications beyond choice, to cross-game similarity and transfer. We formulate specific hypotheses on the types of heuristics players may use in our games, as well as in other games that present the same or similar features: a heuristic that prescribes to choose the strategy with the best risk-return profile, and one that suggests to pick the strategy leading to an attractive and fair outcome for both players.

---

[5] Two referees have pointed out that the suggestion to not use more than 30 seconds to make a decision may have generated an experimenter demand effect pushing subjects to use low-rationality shortcuts. While this is in principle a possibility, we observed that only ten of our subjects never took more than 30 seconds, showing that the suggestion was largely ignored. Moreover, we replicated the experiment in Devetag et al. (2012) using the same payoff matrices and without any recommendation concerning timing: all our findings on the behavior of row players were confirmed in the new experiment.

[6] Although the exchange rate was not announced at the beginning of the experiment, subjects were told the minimum and maximum gain they could obtain.

To sum up, our experimental design is meant to test the following research hypothesis:

Hypothesis 1 ("Feature-Sensitive Strategic Behavior"): a) when the variance of HA is low, strategies FP and HA capture the majority of choices in games with a focal point, and strategy HA captures the majority of choices in games without a focal point; b) when the variance of HA increases, its share decreases, ceteris paribus c) the share of the FP strategy in matrices with focal point is higher than the share of the corresponding strategy in matrices without focal point (XFP strategy).

Hypothesis 2 ("Feature-Based Similarity Perception"): a) a "descriptive feature" has a similar effect in strategically different games by influencing choice behavior in the same direction, and b) keeping all other features fixed, the choice distributions in matrices which are strategically different but similar with respect to the features are closer - statistically - than the choice distributions of matrices which are strategically equivalent but differ with respect to the features.

4. Results and Discussion

4.1 Analysis of aggregate choices

A data overview is given in Figures 2a to 2c, which report the observed frequencies for each row (choice) separately, the 30 games being grouped together. Each figure shows two lines: the continuous lines report choice frequencies in games with focal point, the dashed lines those in games without focal point.[7]

The figure reveals, for each game, marked differences in the choice distributions of the six versions, suggesting that feature manipulation has influenced behavior to a great extent.

Furthermore, some patterns are clear-cut: specifically, the difference in observed frequencies between the same game with and without focal point is evident in most cases, as are the low shares of the EQ strategy (except for the Prisoner's Dilemma) and the effect of increasing the variance of HA. In particular, for each game, differences in the choice distributions of matrices with focal point and low variance (FP_L) and without focal point and high variance (XFP_H) − the two extreme

---

[7] Since the two versions (with and without focal point) of the Weak Link game differ only for the position of the cells in the matrix, to make the comparison more straightforward Figure 2a reports the frequency of row 2 for games WL_FP, and of row 1 for games WL_XFP. Similarly, Figure 2b reports the frequency of row 1 for games WL_FP, and of row 2 for games WL_XFP.

cases − are statistically significant in all games at least at a p level of 0.01, according to a chi-square test.

FIGURE 2a

FIGURE 2b

FIGURE 2c

We now examine Hypothesis 1 in greater detail.

Table 2 summarizes our findings. As hypothesized (Hypothesis 1a), when both features are strong (as in games with focal point and low variance) the corresponding strategies capture the large majority of players' choices, and when the focal point is eliminated, HA increases its attractive power, capturing almost the same shares as in the previous case. The case of DomCol is emblematic, as in DomCol_FP_L only 17% of players choose the equilibrium strategy, even though it is the best response to a column player choosing a strictly dominant strategy, and in DomCol_XFP_L HA is selected by 80% of players.

	FP + HA low var	HA low var (in XFP games)
DomCol	83%	80%
noNE	83%	73%
UniqNE	90%	75%
PD	97%	92%
WL	99%	48% (+48%)

Table 2: Observed frequencies of FP + HA choices in matrices with HA low var and focal point, and frequencies of HA low var choices in matrices without focal point

PD and WL strongly support our hypothesis: less than 5% of players fall outside the FP+HA combination, although in the Prisoner's Dilemma HA = EQ by construction. The only case which apparently contradicts our hypothesis is WL_XFP_L, in which 48% of subjects choose HA and another 48% XFP. However, as previously specified, in the Weak Link the focal point was removed by simply moving the focal cell outside the main diagonal, with no changes in payoffs in order to preserve the game equilibrium structure. This evidence shows that simply moving a focal point away from a central position does not reduce its attractiveness: therefore the frequencies must be interpreted as 96% of players choosing HA+FP, still in line with our hypothesis.

Looking at Table 2, it is noteworthy that the choice pattern in the noNE game is similar to those observed in DomCol and UniqNE, although noNE does not have any pure strategy Nash equilibria. This finding is consistent with a "similarity judgment" approach (Leland, 1994; Rubinstein, 1988), according to which strategy C3 of noNE may be considered as "almost-dominant", since it yields the highest payoff in 2 out of 3 cases, and a not significantly lower payoff in the third case. Since choosing R3 is the best response to a column player choosing an "almost-dominant" strategy, (R3, C3) may be considered a "quasi-equilibrium" in pure strategies (QES, henceforth). This hypothesis is also supported by the behavior of the column players, as their choice distributions in DomCol and noNE are very similar, as shown in Table 3.

Column player	FP (EQ)	XFP (EQ)	Binomial test one-tailed
DomCol HA low	30% (70%)	5% (95%)	0.05
DomCol HA middle	50% (50%)	0% (100%)	0.00
DomCol HA high	35% (65%)	5% (95%)	0.02
noNE HA low	25% (75%)	0% (100%)	0.03
noNE HA middle	45% (55%)	0% (100%)	0.00
noNE HA high	30% (70%)	5% (90%)	0.05
UniqNE HA low	60% (40%)	15% (70%)	0.00
UniqNE HA middle	45% (55%)	30% (70%)	0.26
UniqNE HA high	60% (40%)	25% (70%)	0.03

Table 3: Frequencies of FP and XFP choices for column players, and corresponding p-values. In brackets, frequencies of EQ and QES strategies in the corresponding matrices.

It is reasonable to assume that a certain number of players select the strategy with the highest expected value, assuming, more or less implicitly, that their opponents' choices are equally likely. This behavior is relatively well known for normal form games and has been defined as "Level-1" or "Naive" (Camerer et al. 2004; Costa-Gomes et al., 2001; Stahl and Wilson, 1995). What has not been taken into account so far is the role played by perceived risk in influencing "Level-1" types of reasoning. According to the literature, what matters for "Level-1" players is a strategy-expected value. Instead, in line with previous findings (Warglien et al., 1999), we assume that the attractiveness of the highest expected value strategy is also a function of its safety: therefore, the higher the variance, the lower the attractiveness, ceteris paribus. To the best of our knowledge, no previous study has systematically investigated the role of perceived risk, as measured by payoff variance, in determining the fraction of players who exhibit behavior compatible with a Level 1 type.

Table 4 reports data on the share of HA as a function of variance for the first four games. We discuss the Weak Link case separately.

The table shows that the share of HA always decreases monotonically when the variance of HA increases from low, to medium, to high, except in two cases, in which it stays constant from medium to high (noNE without focal point and PD with focal point). We tested differences between matrices with HA low variance and those with HA high variance by both a chi-square and a binomial one-tailed test. For games DomCol, noNE, and UniqNE, both tests revealed that the differences were statistically significant ($p \leq 0.01$; except in two cases, in which $p \leq 0.5$). Those for the PD without focal point are likewise significant (p-value < 0.01). The PD with focal point is the only case in which the difference is not significant, although the trend is the same as in the other games.

The case of PD is particularly intriguing, since HA corresponds to EQ by construction, and is weakly dominant. Nevertheless, increasing the strategy variance without affecting its dominance induces a shift in behavior. We discuss this finding in detail below.

On average, the frequency of HA passes from 68% (low variance case) to 43% (high variance case).

	HA low variance	HA middle variance	HA high variance	Chi-square test	Binomial test one-tailed
DomCol FP	45%	27%	23%	0.02	0.01
DomCol XFP	80%	48%	43%	0.00	0.00
NoNE FP	52%	37%	20%	0.01	0.00
NoNE XFP	73%	53%	53%	0.00	0.02
UniqNE FP	43%	28%	20%	0.00	0.00
UniqNE XFP	75%	68%	47%	0.00	0.00
PD FP	87%	80%	80%	0.34	0.23
PD XFP	92%	87%	68%	0.00	0.00

Table 4: Frequencies of HA choices for row players, and corresponding p-values obtained by comparing low and high variance frequencies

A different approach must be used for the Weak Link game. Here, the effect of variance cannot be observed directly, but it must be inferred from the share of strategy COS (the strategy giving a constant payoff). Due to equilibrium constraints, while in HA low variance and HA middle variance strategies HA and FP are distinct, in HA high variance two focal points appear: one in the former FP strategy and another in HA. Therefore, instead of testing whether increasing the variance of HA reduces its share, we verify whether it increases the share of COS (the only strategy without focal points). In WL_FP, the frequency of COS passes from 2% in the low variance matrix, to 8% in the medium variance matrix, to 18% in the high variance matrix, whereas in WL_XFP, the frequency rises from 3%, to 12%, to 23%. In both cases, chi-square and binomial tests showed that the differences between low and high variance matrices were statistically significant ($p < 0.01$). We conclude that, in WL too, our hypothesis (1b) is confirmed.[8]

Concerning the role of the focal point in our matrixes (hypothesis 1c), it is noteworthy that the share of FP is always higher (and equal in only one case) than the share of XFP (see Figure 2b). The frequencies of FP, XFP, and the corresponding p-values are listed in Table 5. In the first three game categories − DomCol, noNE and UniqNE − the average difference in share between FP and XFP is 38%. In the case of PD and WL, it falls to 6.5%, and is 25.4% overall.

We made pairwise comparisons of the choice distributions with a chi-square test. The hypothesis was confirmed for games DomCol, noNE and UniqNE and the difference was statistically significant in all 9 comparisons (p-value < 0.01). In PD too, the frequencies of XFP were always smaller than or equal to the corresponding frequencies of FP, but the difference was statistically significant only in the pair with HA medium variance (chi-square test p-value < 0.1; binomial test p-value < 0.5, one-tailed). There are two reasons for this difference: first and most importantly, the focal point in game PD is weaker than in remaining games, due to a lower relative payoff magnitude: consequently, the related strategy is chosen by fewer subjects than in any other game. Second, in the PD, the focal point is eliminated only by breaking the symmetry, with a minimal change in payoff magnitude for the column player and no changes in the payoff of the row player.

In WL too, FP frequencies are higher than those of XFP, although the differences are not statistically significant. As previously said, in the Weak Link, XFP is obtained by simply shifting

---

[8] Analyses of our data performed with the goal of distinguishing the effect of payoff variance (which is the object of our main hypothesis) from the effect of safety (deriving from subjects applying minimax decision rules) were not conclusive. While the two effects are likely to be closely related and hence behaviorally very similar, testing payoff matrices that allow a sharp distinction between the two is an important goal that should be addressed in future research.

the cell position without altering its content. This change apparently does not significantly affect the cell focality[9].

	FP	XFP	P-value chi-square	Binomial test one-tailed
DomCol HA low	38%	2%	0.00	0.00
DomCol HA middle	42%	7%	0.00	0.00
DomCol HA high	43%	5%	0.00	0.00
noNE HA low	32%	7%	0.00	0.00
noNE HA middle	50%	7%	0.00	0.00
noNE HA high	58%	0%	0.00	0.00
UniqNE HA low	47%	13%	0.00	0.00
UniqNE HA middle	45%	3%	0.00	0.00
UniqNE HA high	43%	12%	0.00	0.00
PD HA low	10%	5%	0.58	0.24
PD HA middle	17%	5%	0.07	0.04
PD HA high	10%	10%	0.20	0.50
WL HA low	57%	48%	0.60	0.46
WL HA middle	58%	50%	0.62	0.46
WL HA high	82%	77%	0.73	0.65

Table 5: Frequencies of FP and XFP choices for row players, and corresponding p-values

As regards the importance of the focal point, the behavior of the column players is particularly interesting. The DomCol game presents a strictly dominant strategy for the column player, whereas both noNE and UniqNE present a strategy yielding the highest payoff in 2 out of 3 cells and a slightly lower payoff in the third cell: hence, a large share of FP on the part of column players indicates that its importance is considerable, in view of the available alternatives. The frequencies of FP, XFP, and of the (quasi)-dominant strategies for column players are listed in Table 3. When the focal point is present, 100% of column players choose FP or the EQ(QES) strategy, but very few of them violate strict (or quasi) dominance when the focal point is absent, as shown by the values of the EQ shares shown in brackets. The choice of FP on the part of these players cannot therefore be attributed to error or confusion. Since several strategies have frequency 0, the chi-

---

[9] We remind that the frequency WL HA high variance is obtained by summing the frequencies of FP and HA, since - for structural reasons - two identical focal points appear in that matrix, one for each of these strategies.

square test cannot be applied. We therefore only use the binomial, one-tailed test. The average difference between FP and XFP is 32.8%, and in all but one case it is significant, with p-values of less than 0.05. Altogether, these results confirm our hypothesis 1c and show that, when the difference between FP and XFP outcomes is evident, the effect on subjects' choice behavior is both quantitatively and statistically significant.

Hypothesis 2

Our aim in this study is to show that observed differences in strategies among games sharing the same equilibrium structure follow predictable patterns governed by the presence vs. absence of the descriptive features defined above, which reveal the use of simple decision heuristics on the part of our subjects.

For all our game types, the difference in choice shares between the matrix with all features and that without features is always significant, with a p-value of less than 0.01 (chi-square test). A focal point (according to our definition) is one of these features. We have shown that, even when FP is a strictly dominated strategy, it can still attract a significant fraction of players' choices. This effect was observed in several games, with different equilibrium structures, both symmetric and non-symmetric.

Another feature that influences strategic behavior is the strategy giving the highest average payoff (HA) when it is perceived as a “safe” option (low variance). In this case too, HA determines similar effects in different games, and the importance of the “safety” attribute is revealed by the emergence of an inverse relationship between the share of players choosing HA and its variance level.

Altogether, our results show that some features affect behavior in the same direction, regardless of the game-theoretical properties of the strategic situation at hand. Therefore, it may be hypothesized that strategically different games are perceived as similar when they share some or all of these features. Furthermore, we propose that games sharing the same features generate choice distributions that are so similar as to be statistically indistinguishable.

Table 6 lists p-values obtained by comparing games with the same features and different strategic structures. We omit the Weak Link because, comparison-wise, its strategic structure is too different. The data show that, for the game types DomCol, noNE and UniqNE, in most of the comparisons, frequency distributions in games sharing the same features do not appear to be significantly different.

Hence, whereas frequencies differ significantly when the same game type is compared with and without features (as shown in the previous hypothesis), when the latter remain unaltered but the game structure changes, players' strategic behavior remains largely invariant, suggesting that the difference is not perceived as such in the aggregate.

In further support to our hypothesis, it must be noted that the frequencies of DomCol, noNE and UniqNE are all significantly different (according to a chi-square test) from one another only in the XFP high var case, when all features are removed and hence the underlying equilibrium structure may result more clearly visible.

		Chi-square test			**Binomial test, two-tailed HA/no HA**			**Binomial test, two-tailed FP/no FP**		
		noNE	**UniqNE**	**PD**	**noNE**	**UniqNE**	**PD**	**noNE**	**UniqNE**	**PD**
HA low var FP	**DomCol**	0.72	0.47	0.00	0.58	1.00	0.00	0.57	0.46	0.00
	noNE		0.21	0.00		0.46	0.00		0.13	0.01
	UniqNE			0.00			0.00			0.00
HA middle var FP	**DomCol**	0.05	0.83	0.00	0.01	1.00	0.00	0.46	0.85	0.00
	noNE		0.18	0.00		0.44	0.00		0.71	0.00
	UniqNE			0.00			0.00			0.00
HA high var FP	**DomCol**	0.23	0.88	0.00	0.82	0.82	0.00	0.14	1.00	0.00
	noNE		0.16	0.00		1.00	0.00		0.14	0.00
	UniqNE			0.00			0.00			0.00
HA low var XFP	**DomCol**	0.36	0.04	0.02	0.52	0.66	0.12	0.36	0.04	0.61
	noNE		0.26	0.01		1.00	0.02		0.36	1.00
	UniqNE			0.05			0.03			0.21
HA middle var XFP	**DomCol**	0.85	0.08	0.00	0.71	0.04	0.00	1.00	0.68	1.00
	noNE		0.23	0.00		0.13	0.00		0.68	1.00
	UniqNE			0.02			0.03			1.00
HA high var XFP	**DomCol**	0.03	0.07	0.00	0.04	0.19	0.00	0.24	0.32	0.49
	noNE		0.02	0.00		0.58	0.13		0.02	0.04
	UniqNE			0.04			0.03			1.00

Table 6: Comparison of games with same key features and different strategic structures.

4.1.1 Unpacking focality

Our notion of focal point and its properties extends well beyond the domain of equilibrium outcomes in (symmetric) coordination games. It has already been shown that the share of the FP strategy is always higher than that of XFP, but in this section we try to identify the main sources of focality by investigating why some of the differences are considerably more remarkable than others.

There are 4 attributes of a game outcome which, intuitively, may be judged to be relevant in determining focality:

1. payoff magnitude (“significantly” greater than most of the other payoffs)
2. symmetry of payoffs
3. centrality of the cell (or positioned in the main diagonal in WL)
4. Pareto-efficiency

Payoff magnitude refers to the magnitude of a cell payoff, when compared with the other payoffs the same player can get elsewhere in the matrix. To be attractive, a payoff has to be higher than most other payoffs, although not necessarily the highest (which would be chosen by “maximaxi” players). For example, in DomCol_FP_L, the payoff of the focal point is “significantly” higher than the other payoffs, giving 80 ECUs (Experimental Currency Units) against 40 of the second-highest payoff. Conversely, in PD, the payoff of the focal point is not significantly higher, as in PD_FP_L there are 4 other cells which can give the row player the same payoff as the focal cell (35 ECUs).

Symmetry of payoffs indicates that the payoffs of the two players are identical.

Centrality of the cell refers to the position of the cell in the matrix. Motivated by the results of previous experiments (Warglien et al., 1999). the focal point was always located at the center of the matrix, except in the WL, where (due to the presence of three symmetric cells with increasing magnitude) the symmetric cells were positioned on the main diagonal in order of decreasing payoff magnitude.

The choice of Pareto-efficiency as an attribute instead of “Nash Equilibrium” differentiates our definition of a focal point from previous definitions used in the literature. We assume that players do not initially reason strategically in a game-theoretical sense: therefore, we consider that it is more important for the attractiveness of an outcome to be Pareto-efficient rather than an equilibrium.

A focal point is an outcome (a cell) and not a strategy. Since only choices of strategies are observed and motivations for choices are not observed, the strategies yielding a focal point were built in such a way that outcomes other than the focal point look particularly unattractive. In all games, one of the two remaining cells gives the lowest possible payoff to row players, and in all games except the WL the remaining cell yields the second lowest payoff. In addition, one of these two cells gives the highest possible payoff to column players; hence, subjects should avoid picking FP if they imagine that column players might go for their highest payoffs (which in our games coincides with the equilibrium strategy for column players).

In these games, two types of focal points can be found on the basis of the four attributes listed. The first is the focal point for games DomCol, noNE, UniqNE, and WL, which satisfies the attributes of payoff magnitude, symmetry of payoffs, centrality of the cell, and Pareto-efficiency. The second is the focal point for the PD, which satisfies symmetry of payoffs, centrality of the cell, and Pareto-efficiency, but not payoff magnitude.

We removed focal points in one of three different ways: first, as in game types DomCol, noNE, and UniqNE, by breaking the symmetry of payoffs and by reducing their magnitude, so that the cell that used to be focal satisfies only the attribute of centrality and Pareto-efficiency. Second, as in the Weak Link, by simply shifting strategy positions so as to place all the cells with symmetric payoffs outside the main diagonal. Therefore, in the Weak Link, the XFP outcome satisfies the attributes of payoff magnitude, symmetry of payoffs, and Pareto-efficiency. Third, as in the PD, by simply reducing the payoff of the column player. Since both payoffs were already relatively small, the payoff decrease in this case is slight. Here XFP satisfies centrality of the cell and Pareto-efficiency (in 2 out of 3 matrices).

Table 7 lists attributes and choice shares for a sample of payoff matrices. The data clearly suggest that some of the attributes are an important source of focality whereas others are not.

Type of focal point	PD_L		DomCol_M		WL_L		PD_XFP_L	DomCol_XFP_L
Strategy	FP	XFP	FP	XFP	FP	XFP	DOM	XFP
Payoff magnitude			X		X	X		X
Symmetry of payoff	X		X		X	X	X	
Centrality ofcell	X	X	X	X	X			X
Pareto efficiency	X	X	X	X	X	X		X
Frequency	10%	5%	42%	7%	57%	48%	3%	2%

Table 7: Attributes and choice frequencies for a sample of cells

Let us first analyze PD_FP_L, in which the FP strategy is not particularly successful, being chosen only by 10% of players. As the difference with PD_XFP_L is not significant, we infer that the joint presence of symmetry of payoffs, centrality of the cell, and Pareto-efficiency is not sufficient to trigger focality.

We then analyze game DomCol, that represents also games noNE, and UniqNE, since their FP and XFP cells share the same attributes. The FP strategy in these games is highly attractive, reaching a share ranging from 32% to 47% in the low variance case. In addition, in all versions, the differences between FP and XFP are always significant, suggesting that symmetry of payoffs and payoff magnitude (the attributes removed in XFP) are a key source of focality. Instead, since XFP is rarely selected, it appears that Pareto-efficiency and centrality of the cell are two attributes of minor or no importance, as already indicated by data related to the PD.

In the Weak Link, FP has the strongest attractive power and, when matrices with the same features are compared, it obtains the highest share. Although the share of FP is always higher than that of XFP in the Weak Link, the difference is never significant, again indicating that centrality of the cell plays a minor role in determining focality.

Lastly, we consider the separate effects of symmetry of payoffs and payoff magnitude: although the two attributes show considerable attractive power when together, neither seems to create a focal point when alone. In PD_XFP_L, only 3% of subjects chose strategy DOM[10], although it contains a symmetric cell yielding an "acceptable" gain to both players. Similarly, in DomCol_XFP_L, only 2% of row players chose strategy XFP, which yields the highest (although not symmetric) gain compared with other matrix cells.

Altogether, these results suggest that cell focality in a non-symmetric game is mainly due to the joint effect of payoff magnitude and symmetry of payoffs, whereas centrality of the cell and Pareto-efficiency play a minor role. The two attributes, when present in isolation, lose much of their attractive power. This finding is consistent with the results of Biel (2009), in which introducing cells with symmetric payoffs in normal form games turned out to be irrelevant in modifying players' strategic behavior.

So far we have stated that the attractiveness of a focal point is due to its structure, meaning that its features make it a "natural" cooperative choice in the absence of communication or feedback. An alternative explanation may be that the focal point is chosen being the outcome that yields the highest payoff sum. Fairness-based explanations of out-of-equilibrium play are widespread, and behavioral models such as that of Costa-Gomes et al. (2001) include an "Altruistic" type, who systematically opts for the cell with the highest payoff sum. In order to test whether players select FP for this reason, in the following we analyze the relative attractiveness of the "fair" cell, defined as the one with the highest payoff sum.

In matrices with a focal point, the focal point is always the fair cell. In PD_FP_L, PD_FP_H and WL_FP_H, another cell yields the same payoff sum as the focal point (in strategies EQ/HA, EQ/HA, and HA). In all matrices with a focal point, the strategies corresponding to the fair outcomes are

---

[10] DOM is a weakly dominated strategy, present in our modified Prisoner's Dilemma

chosen by a share ranging from 32% to 87%. The only exception is PD_FP_M, in which the strategy leading to the only fair cell – FP – is only chosen by 17% of subjects, preliminary evidence of the scarce importance of payoff sum as a choice criterion.

Let us now examine fair cells in matrices without focal point. The cases of PD and WL are not informative: in PD, fair cells are always selected by the EQ/HA strategy, and another fair cell appears in XFP as well in PD_XFP_M and PD_XFP_H. In the WL, the focal point is not really removed, but it is only shifted to a different position and this change does not affect its salience. We therefore analyze the case of games DomCol, noNE, and UniqNE.

Here, the cell that used to be focal is still the fair cell in 8 out of 9 matrices, but the share of the corresponding strategy (XFP) ranges from 0% to a maximum of 7%, whereas in matrices with focal point the share of the associated strategy ranges from 32% to 58%. We interpret this difference as strongly supporting the hypothesis that the attractiveness of a focal point is not related to its being the cell with the highest payoff sum, but to the attributes already singled out.

Finally, further evidence of the scarce importance of fairness-driven preferences for our focal points comes from applying the model in Costa-Gomes et al. (2001) to our data: only 4.6% of our subjects is categorized as "Altruistic"(see section 4.2).

Concluding, the symmetry and magnitude of payoffs seem to make the focal cell an "obvious" choice for both players, triggering spontaneous coordination. Clearly, payoff symmetry makes the focal point a fair outcome by definition (as is the result of applying the "equality rule" which Mehta et al. (1994) find as the most frequently used in a series of assignment games), but we argue that subjects select it for reasons which have to do with Schelling salience (Bacharach, 1999; Mehta et al., 1994; Sugden, 1993): that is, out of cognitive processes akin to those which are thought to be triggered by equilibrium focal points in games of coordination.

The last consideration concerns theories of choice in games based on collective preferences such as team reasoning. Team reasoning (Bacharach 1999, 2006; Gold and Sugden 2007; Sugden 1993) has been defined as a decision criterion based on collective rather than individual rationality. A player asks herself not "what do I want and what should I do to achieve it", but rather "What do we want, and what should I do to help achieve it" (Colman et al. 2008, p. 389). It follows that in games with Pareto-ranked equilibria, team reasoning prescribes that players should select the payoff-dominant equilibrium, and in games other than coordination games, team reasoning prescribes to select strategies leading to the Pareto-efficient outcome (even if it is not an equilibrium); since we have shown that the attribute of Pareto-efficiency, when present alone, does not attract a significant share of preferences in our games, we can conclude that team reasoning is not able to explain our data.

Team reasoning also fares worse in our games compared to other 3x3 games with Pareto-efficient, non-equilibrium outcomes in which it has been tested against equilibrium predictions (Colman et al. 2008); in this study, the share of the strategy leading to the Pareto-efficient outcome is always the highest and always higher than 50 per cent, thus predicting behavior substantially more accurately than Nash equilibrium. The main difference between the two sets of matrices is that in Colman et al. the Pareto-efficient outcome is always symmetric (i.e., a focal point, based on our definition). Therefore, by comparing our results with their results we can obtain further indirect support to the idea that payoff symmetry is a fundamental component of a focal point.

4.2 Analysis of individual behavior

In this section we investigate the effect of features on individual behavior by employing three different approaches. First we apply the model in Costa-Gomes et al. (2001) to our data. We then apply the Cognitive Hierarchy model (Camerer et al. 2004), to verify if and to what extent, subjects' levels of reasoning are affected by features manipulations according to this model. Lastly, we analyze correlations among several variables.

Costa-Gomes et al. (2001) identify nine types of strategic behavior, which summarize a wide range of possible decision rules a player can apply in a game: Altruistic (an agent aiming at the cell that maximizes the sum of his own and his opponent's payoff), Pessimistic (an agent choosing the strategy with the highest minimum payoff), Naïve (an agent picking the strategy with the highest average payoff, under the assumption that the opponent's choices are equally likely), Optimistic (a player aiming at the highest payoff for herself), L2 (an agent who best responds to a Naïve opponent), D1 (an agent who is able to single out a dominated strategy to then assign equal probability to the remaining choices of her opponent), D2 (an agent performing two rounds of iterated elimination of dominated strategies), Equilibrium (an agent who selects equilibrium strategies), Sophisticated (an agent who best responds to the probability distribution of his opponent's decisions). In Costa-Gomes et al., of the nine types, only eight are actually used, since two of them (Naïve and Optimistic) coincide in all the games tested. The article presents a mixture model that assumes a specific distribution of types, and assigns a corresponding probability of error to each type (trembling hand). The distribution of types, as well as the probability of error associated with each type is estimated according to a maximum likelihood, error-rate method.

The parameters estimated based on subjects' choices are reported in Table 8.

Subjects in Costa-Gomes et al. appear to be fairly refined in their reasoning patterns, since the majority (all except the 20% acting naively) seem able to recognize the game strategic structure or at least some of its most obvious elements (e.g., through the recognition of a dominant strategy, or through best responding to possible moves by the opponent). Nonetheless, subjects seem to not pay sufficient attention to their opponent's behavior, given that none of them is classified as Sophisticated.

In order to apply the model to our data, we proceeded by classifying each choice according to the pre-defined set of strategic types defined in Costa-Gomes et al. (2001). Of the thirty payoff matrices we used, seventeen had to be excluded because more than one strategy resulted consistent with the same strategic type. Although using all payoff matrices would have generated a more refined estimation, our restricted sample size is still in line with that used in other studies (see for example Stahl and Wilson, 1995).

Of the nine types presented in Costa-Gomes et al., only five could be applied to our games. Types D1, D2, and L2 always coincided with Equilibrium, therefore we simply created the Equilibrium type, which includes them all. Pessimistic and Naïve coincided as well, and we labeled the corresponding type Pessimistic/Naïve. The five resulting strategic types are then: Altruistic, Pessimistic/Naïve, Optimistic, Equilibrium, and Sophisticated. For each type, the corresponding strategy could be identified in each matrix.

The results of the estimations are reported in Table 8.

		Altruistic	Pessimistic	Naive	Optimistic	Equilibrium	L2	D1	D2	Sophisticated
Costa-Gomes et al., 2001	Estimated frequency	0.000	0.000	0.199*	*	0.160	0.344	0.298	0.000	0.000
	Error rates	-	-	0.285*	*	0.165	0.233	0.276	-	-
Our study	Estimated frequency	0.046	0.286°	°	0.000	0.142*	*	*	*	0.524
	Error rates	0.490	0.332°	°	-	0.410*	*	*	*	0.614

Table 8: Parameters estimated in Costa-Gomes et al. (2001) and with our data

The high error rates suggest that these types are not capturing our data well as Costa-Gomes et al.'s. Nonetheless, some interesting observations can be made.

The shares obtained in our games differ largely from those observed in their experiment. According to their findings we should have observed around 80 percent of Equilibrium, no Altruistic, no Sophisticated, and 20 per cent of the remaining two types combined. This striking difference can be explained by taking into account the effect of the game features and the use of the corresponding decision heuristics.

More specifically, the large share of Pessimistic/Naïve can be explained by looking at the game features. The Pessimistic/Naïve type always selects the strategy with the highest average payoff, corresponding in our matrices to strategy HA, yielding the highest average payoff (Naïve) and being the minimax strategy at the same time (Pessimistic).

Type Equilibrium is less frequent than in Gosta-Gomes et al. (2001), since in our games the equilibrium strategy was a dominant strategy for the row player only in two out of thirteen matrices (corresponding to the PD), as opposed to their experiment in which a large part of the games had a strictly dominant strategy. Note, incidentally, that whenever the equilibrium strategy is strictly dominant, the Equilibrium type by definition coincides with Naïve. More generally, it may be argued – and could be object of future research – that whenever subjects select a dominant strategy (and hence behave consistently with equilibrium), they do so not necessarily because they recognize the dominance relation, but because they select the strategy with the highest expected value. Note also that in the *Prisoner's Dilemma*, our subjects move away from HA (even if dominant) when its variance is high, giving further support to this conjecture. Therefore, previous models may have overestimated the percentage of Equilibrium types whenever a dominant strategy is present[11].

The Altruistic and Optimistic types are the only types prescribing the selection of a Focal Point strategy. Altruistic always selects the Focal Point strategy, while Optimistic does so in 2 of 3 cases. Since in the games used in Costa-Gomes et al. no focal points were available, the difference in share of the Altruistic type seems to corroborate the hypothesis that Focal Points are attractive.

The most surprising result is the large share of Sophisticated subjects. While in Costa-Gomes et al. no subjects were categorized under the Sophisticated type, more than half of our subjects belong to this group. We can try to analyze their behavior in greater detail to infer the reasons for this high percentage. Of the 13 games taken into consideration, in 5 the best reply is HA, in 5 FP, and in the remaining 3 EQ. Of the 13 matrices considered, 7 have a Focal Point, and 5 of these have a Focal Point as best reply. In addition, in none of the matrices considered XFP (a strategy with no attractive power according to us) is a best reply.

---

[11] A study that indirectly supports this conjecture is Andreoni (1995). He disentangles kindness from confusion as possible sources of cooperation in public goods games, showing that they are both present and important. While labeling the choice of a strategy based on its risk/return profile "confusion" is highly questionable, nevertheless it is a choice strategy leading to cooperative behavior that does not derive from kindness.

Let us first analyze the 5 matrices in which HA is a best reply. In none of these (matrix n. 1, 16, 17, 20, 22 in Table 1) HA has a high variance. In fact, three have low variance and two have medium variance. In addition, only one of the five matrices includes a strong Focal Point. Three lack a Focal Point and one has a “weak” Focal Point. Column players in the five matrices choose EQ/HA with a percentage ranging from 70% to 100%. Therefore, behavior of Sophisticated players in these matrices is observationally equivalent to the choice of HA (possibly out of risk-return considerations) on both sides.

Let us then consider the five matrices with FP as best reply. By definition, all these matrices include a Focal Point (matrices n. 2, 3, 13, 14, 15); in addition, in all matrices the FP strategy for the column players is dominated or quasi-dominated. Nevertheless, the percentage of column players picking the FP strategy instead of the equilibrium strategy ranges from 35% to 60%. Hence, Sophisticated players in these matrices select the FP strategy expecting the other players to do the same, and being right in their expectations in roughly 50% of the cases. This behavior is on average what we expect to find in the presence of a Focal Point, which can be defined as such if both players recognize focality and expect the other player to recognize it as well.

Finally, the remaining three matrices in which the equilibrium strategy is the best reply (matrices n. 4, 5, and 6) are all without Focal Point, with a percentage of column players picking the Equilibrium strategy ranging from 95% to 100%.

Concluding this section, notwithstanding the necessarily limited data set, the findings provide support to our hypothesis that subjects adapt their behavior to the game they face, and that taking into consideration heuristic-based behavior can help explain the main difference we find in the distribution of player types in the two samples.

We then tried to estimate the average number of reasoning steps that our subjects performed by applying the Cognitive Hierarchy Model. The Cognitive Hierarchy model (Camerer et al., 2004) is a level-k model that assumes types follow a Poisson distribution. As any other level-k model, subjects are divided into different strategic categories according to their level of sophistication. Each subject assumes to be more sophisticated than the others, and chooses his strategy as the best response to a distribution of opponents ranging from level 0 to level k-1 (where k is the level of sophistication of the subject himself). The single model parameter τ corresponds to the average reasoning level of the subjects’ sample.

We estimated the parameter values for each game by choosing the value of τ so as to minimize the mean square deviation (MSD) between the observed and the estimated frequency.

Table 9 shows the value of τ averaged across features and separately by game type. The values of τ differ considerably across games, ranging from a minimum of 0 (indicating that subjects' reasoning steps correspond to a random choice) to a maximum of 3.32 (the average subject performs 3 steps of reasoning). More "intuitive" games (like the Weak Link, with an average parameter equal to 0.33) are apparently solved without reasoning strategically about the opponents' possible moves in contrast to games in which a "preferable choice" for the opponent can be more easily singled out (like the *Prisoner's Dilemma*, average parameter = 1.41).

Interestingly, the average parameter in games with focal point equals 0.72, while it doubles (1.43) in games without focal point. Focal points, in line with our hypothesis, seem to trigger more intuitive responses (as in Kuo et al. 2009). Regarding HA variance, no notable differences seem to emerge on average. The value of τ in games with HA low variance in fact, almost equals that in games with HA high variance (1.14 in the former versus 1.23 in the latter). However, by comparing, again, the extreme cases within each game class (i.e., matrices with Focal Point and HA low variance vs. matrices without Focal Point and HA high variance) it can be noticed that in three cases out of five the difference in the value of τ goes in the direction hypothesized, and in two cases remarkably so (0.19 vs. 1.9 in the DomCol game class, and 0.32 vs. 3.08 in the noNe game class).

Game	FP/XFP	Av. Parameter	HA_L	HA_M	HA_H	Av. τ FP/XFP
DomCol	FP	0.83	0.19	0.00	0.00	0.06
	XFP		1.64	1.23	1.90	1.59
noNe	FP	1.40	0.32	0.05	0.00	0.12
	XFP		1.64	3.32	3.08	2.68
UniqNe	FP	1.42	2.56	0.00	2.74	1.77
	XFP		0.98	1.11	1.12	1.07
PD	FP	1.41	1.61	1.20	1.20	1.34
	XFP		2.08	1.61	0.74	1.48
WL	FP	0.33	0.13	0.00	0.86	0.33
	XFP		0.25	0.08	0.64	0.32
		Av. τ HA	1.14	0.86	1.23	
		Av. τ FP	0.72			
		Av. τ XFP	1.43			

Table 9: Cognitive Hierarchy parameter values estimated for each game type and feature

We state that heuristic-based behavior could be a promising avenue to extend CH models. Camerer et al. (2004), in fact, openly suggest that future research on CH models should try to endogenize the mean number of thinking steps, presumably out of a cost-benefit analysis by which players weight the marginal benefit of further reasoning effort against cognitive constraints. We suggest that the use of heuristics triggered by the presence of attractive features may enter such calculus and be weighed against the use of more "rational" choice algorithms, helping to explain the great variability usually observed in the value of τ across games.

Our analysis so far has revealed that a substantial proportion of subjects exhibit choice patterns that seem at odds with orthodox strategic reasoning. However, are subjects really non-strategic? We can judge how different subjects perform by calculating their `strategic IQ'. Following Bhatt and Camerer (2005), we calculated each subject's expected earnings by matching his choice in every matrix with the population average of all the column players. This quantity measures a subject's strategic IQ. In order to test whether different choices were more "strategic" than others, we calculated the correlation between an individual's strategic IQ and the number of HA/FP/EQ choices she made in the game. Furthermore, we can ask whether part of our findings can be explained by subjects' differing attitudes to risk. For this reason, we calculated the correlation between risk aversion, as measured by the score obtained in the H&L test, and choices in the games. Table 10 reports the results of the correlation test[12].

	H&L	Strategic_iq
HA	-0.062	-0.503**
FP	0.078	0.284*
EQ	0.147	0.466**
Strategic_iq	0.215	-

Table 10: correlation coefficients, Spearman correlation test. ** correlations with p-values < 0.01, * correlations with p-values < 0.05.

Some observations are noteworthy: first, strategic IQ negatively correlates with the choice of HA, suggesting, once again, that players selecting HA do so out of a choice process that seems to ignore the opponents' incentives and motivations, and this gets reflected in their relatively poorer strategic performance. On the contrary, strategic IQ is positively correlated with the choice of a focal point, indicating that choices of focal points are perfectly compatible with a strategic approach, and that

---

[12] We obtain the same results with a Pearson correlation test.

players who select focal points do so on the expectation that other players will select them as well. Finally, strategic IQ is also positively correlated with equilibrium choices. These results do not change if we consider spurious combinations such as HA/EQ, etc.

Risk aversion is positively correlated with equilibrium choices and with strategic IQ, whereas no correlation emerges between risk attitudes and choices of HA and FP. No correlation was also found between risk attitudes and variance levels of HA.

4.3 Analysis of response times

To gain further insight into subjects' choice processes, we finally analyze differences in response times.
Figure 3 shows average response times, disaggregated by game class and matrix version.

FIGURE 3

Some recent studies of gaming behavior employ response time as a means to explore subjects' decision-making processes, as opposed to the more invasive and expensive methods based on study of neural activity. Both Rubinsten (2007) and Piovesan and Wengström (2009) analyze the relationship between response times and social preferences. Rubinstein's study finds that fair decisions take a shorter response time than egoistic (more rational) ones, whereas Piovesan and Wengström (2009) seem to find the opposite, although the two experimental designs differ in many respects. In a recent fMRI study on gaming behavior, Kuo et al. (2009) found that subjects took a much longer time, on average, to choose a strategy in dominance-solvable games than in coordination games, and different areas of the brain were activated when players faced instances of the two classes of games. According to these findings, the authors suggested the existence of two different "strategizing" systems in the brain, one based on analytical reasoning and deliberation and the other on intuition and a "meeting of the minds".

As proposed by Kuo et al. (2009), we also hypothesize that matrices with a focal point trigger intuitive reasoning and hence require a shorter response time than matrices without a focal point, which are presumed to activate analytical reasoning.

We do not expect the relation between response time and type of game to be as notable as reported by Kuo et al. (2009), as the two game types in their study were indeed strategically different, whereas in our case they only differ in the presence of a focal point, as defined earlier.

Nonetheless, the individual response time for matrices with a focal point is significantly shorter than that for matrices without focal point, according to a paired t-test ($p < 0.01$, two-tailed[13]). Hence, our data support the hypothesis that matrices without focal point require more cognitive effort. Note that the significance of results holds, although some subjects did not select the focal point strategy in the matrices which contained it, and those who did not presumably employed the same type of analytical reasoning used for games without focal point.

The second important finding is the increased response time which can be observed when the variance of the HA strategy increases (from low, to medium, to high). The increasing pattern is clear-cut in Figures 3 and 4, which shows average response time when games are aggregated according to variance level. The figures show that increasing the variance leads to large increases in response time.

Response time averages 17.71 in the low variance case, 20.98 in the middle variance case, and 23.66 in the high variance case. Pairwise differences of individual response time are significant according to a paired t-test, two-tailed ($p = 0$ for all cases: low var-middle var, low var-high var, and middle var-high var[14]).

We then compared the two "extreme" cases according to these findings, i.e., matrices with focal point and low variance - which should be the fastest to process - and matrices without focal point and with high variance - which should instead require the highest cognitive effort. The difference in response time was indeed remarkable, increasing on average from 17.61 to 24.27 from the first to the second groups. Also in this case, the differences in individual response time were significant (paired t-test test, $p = 0$, two-tailed).

FIGURE 4

No significant correlations were found between individual response times and either number of FP choices or number of HA choices. Instead, a significant correlation was found between individual response times and number of EQ choices. The correlation coefficient is positive and is .273

---

[13] The same result was obtained by a non-parametric Wilcoxon signed rank test ($p < 0.01$, two-tailed).

[14] The same result was obtained by a non-parametric Wilcoxon signed rank test ($p = 0$, two-tailed).

(Spearman's rho coeff., p = 0.035, two-tailed) when choices from the modified PD (in which EQ = HA) are included, and is .331 (Spearman's rho coeff., p = 0.01, two-tailed) when choices from modified PD are excluded, leaving only “pure” EQ choices.

This finding shows that the players who were more likely to choose the equilibrium strategy took longer to respond, as found by Kuo et al. (2009). These correlation results also indicate that choices of FP or HA generally derive from the use of intuitive heuristics, rather than from beliefs in other players' irrationality. In fact, if the latter were the case, i.e., if players always correctly identified the equilibrium strategy even when they did not select it, we would not observe higher response times for EQ choosers.

5 Conclusions

In this paper we show that behavior in one-shot games may be explained by a set of very simple behavioral rules which eschew optimization and are triggered by the presence of salient features: two of such features are a “focal point” and a strategy with high expected value and low variance.

More specifically, we show that the attractive power of focal points extends to asymmetric games and non-equilibrium outcomes, and identify two attributes (payoff symmetry and payoff magnitude) which, when jointly present, are the two factors most frequently responsible for making an outcome focal.

We also show that the presence of a strategy with high expected value and low variance (a “safe”, attractive strategy) is a strong choice attractor.

Together, the strategy yielding the focal point and the safe strategy explain most of players’ choices. Subjects react in similar ways to games with the same features, regardless of their game-theoretical category, and treat formally equivalent games differently when they differ with respect to descriptive features.

We suggest how our findings may be used to extend the predictive power of level-k and CH models. Finally, analysis of response times shows that matrices with focal points are faster to process than matrices without them, and that there is a direct relationship between the variance level of the HA strategy and average response times. Equilibrium choices take longer than other choices, indicating that out-of-equilibrium choices are not due to beliefs in other players' irrationality, but rather to the use of low rationality heuristics based on simplified mental representations of the strategic situation to hand (Devetag and Warglien, 2008, Weizsäcker 2003).

Future research will have to proceed in two complementary directions: exploring subjects' decision-making processes and similarity perceptions in greater depth, through the use of eye-tracking techniques and the elicitation of direct similarity judgments; and developing a comprehensive theory of cross-game similarity, based on experimental results which may help to model and predict cross-game transfer and generalization.

Acknowledgments

We thank Christopher Gilbert, Paola Manzini, Marco Piovesan, Ondrej Rydval, Theodore L. Turocy, Enrico Zaninotto, participants at the IMEBE 2010 in Bilbao, and seminar participants in Trento for their useful comments and suggestions. We also thank Franz Dietrich, Ido Erev, Hykel Hosni, Ariel Rubinstein, Karl Schlag and other participants at the Workshop on Rationality, Heuristics and Motivation in Decision Making, Scuola Normale Superiore, Pisa, Italy, November 12-14 2010. We thank the staff of the CEEL experimental laboratory for help in implementing experiments. We gratefully acknowledge financial support from the R.O.C.K. research group, University of Trento. The funding source had no role in the design of the present experiment, in the collection, analysis and interpretation of data, in the writing of the report, and in the decision to submit the paper for publication. The usual caveats apply.

References

Andreoni, J., 1995. Cooperation in Public-Goods Experiments: Kindness or Confusion? Amer. Econ. Rev. 85(4), 891-904.

Bacharach, M., 1999. Interactive Team Reasoning: A Contribution to the Theory of Co-Operation. Res. Econ. 53(2), 117-147.

Bacharach, M., 2006. In N. Gold and R. Sugden (Eds.), Beyond individual choice: Teams and frames in game theory. Princeton, NJ: Princeton University Press.

Bhatt, M., Camerer, C. F., 2005. Self-Referential Thinking and Equilibrium as States of Mind in Games: fMRI Evidence. Games Econ. Behav. 52(2), 424-459.

Biel, P. R., 2009. Equilibrium Play and Best Response to (Stated) Beliefs in Normal Form Games. Games Econ. Behav. 65(2) 572-585.

Bosch-Domènech, A., Montalvo, J. G., Nagel, R., Santorra, A., 2002. One, Two, (Three), Infinity, … : Newspaper and Lab Beauty-Contest Experiments, Amer. Econ. Rev. 92(5), 1687-1701.

Bosch-Domènech, A., Vriend, N. J., 2008. On the Role of Non-Equilibrium Focal Points as Coordination Devices (Unpublished Results, Working Paper, Queen Mary University of London, No. 621).

Camerer, C. F., 2003. Behavioral Game Theory: Experiments in Strategic Interaction. Princeton: Princeton University Press.

Camerer, C. F., Ho, T., Chong, J., 2004. A Cognitive Hierarchy Model of Games. Quart. J. Econ. 119(3), 861-898.

Colman, A. M., Pulford B.D., Rose J., 2008. Collective rationality in interactive decisions: Evidence for team reasoning, Acta Psychologica, 128, 387-397.

Costa-Gomes, M., Crawford, V. P., 2006. Cognition and Behavior in Two-Person Guessing Games: An Experimental Study. Amer. Econ. Rev. 96(5), 1737-1768.

Costa-Gomes, M., Crawford, V. P., Broseta, B., 2001. Cognition and Behavior in Normal-Forma Games: an Experimental Study. Econometrica 69(5), 1193-1235.

Costa-Gomes, M., Weizsäcker, G., 2008. Stated Beliefs and Play in Normal-Form Games. Rev. Econ. Stud. 75(3), 729-762.

Crawford, Vincent P., 2003. Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions. Amer. Econ. Rev. 93(1), 133-149.

Crawford, V. P., Gneezy, U., Rottenstreich, Y., 2008. The Power of Focal Points Is Limited: Even Minute Payoff Asymmetry May Yield Large Coordination Failures. Amer. Econ. Rev. 98(4), 1443-1458.

Crawford, V. P., Iriberri, N., 2007. Level-k Auctions: Can a Non-Equilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?. Econometrica 75(6), 1721-1770.

Devetag, G., Di Guida, S., Polonio, L., 2012. An Eye-tracking Study of Feature-based Choice in One-shot Games (Unpublished Results).

Devetag, G., Warglien, M., 2008. Playing the Wrong Game: An Experimental Analysis of Relational Complexity and Strategic Misrepresentation. Games Econ. Behav. 62(2), 364-382.

Fischbacher,, U., 2007. Z-Tree: Zurich Toolbox for Ready-made Economic Experiments. Exper. Econ. 10(2), 171-178.

Gilboa, I., Schmeidler, D., 1995. Case-Based Decision Theory. Quart. J. Econ. 110(3), 605-639.

Goeree, J. K., Holt, C. A., 2011. Ten Little Treasures of Game Theory and Ten Intuitive Contradictions. Amer. Econ. Rev. 91(5), 1402-1422.

Gold, N. and Sugden, R., 2007. Theories of Team Agency. In Peter, F. and Schmid, H. B. (Eds.) Rationality and Commitment, 280–312. Oxford: Oxford University Press.

Ho, T. H., Camerer, C., Weigelt, K., 1998. Iterated Dominance and Iterated Best Response in Experimental "p-Beauty Contests". Amer. Econ. Rev. 88(4), 947-969.

Holt, C. A., Laury, S. K., 2002. Risk Aversion and Incentive Effects. Amer. Econ. Rev. 92(5), 1644-1655.

Jehiel, P., 2005. Analogy-Based Expectation Equilibrium. J. Econ. Theory 123(2), 81–104.

Kahneman, D., Tversky, A., 1979. Prospect Theory: An Analysis of Decision Under Risk . Econometrica 47(2), 263-292.

Knez, M., Camerer, C., 2000. Increasing Cooperation in Prisoner's Dilemmas by Establishing a Precedent of Efficiency in Coordination Games. Organizational Behavior and Human Decision Processes 82(2), 194-216.

Kuo, W., Sjöström, T., Chen, Y., Wang, Y., Huang, C., 2009. Intuition and Deliberation: Two Systems for Strategizing in the Brain. Science 324(5926), 519-522.

Leland, J. W., 1994. Generalized Similarity Judgments: An Alternative Explanation for Choice Anomalies. J. Risk Uncertainty 9(2), 151-172.

McKelvey, Richard D., McLennan, A. M., Turocy, T. L., 2010. Gambit: Software Tools for Game Theory, Version 0.2010.09.01. http://www.gambit-project.org.

McKelvey, R. D., & Palfrey, T. R., 1995. Quantal Response Equilibrium for Normal Form Games. Games Econ. Behav. 10(1), 6-38.

Mehta, J., Starmer, C., Sugden, R., 1994. The Nature of Salience: An Experimental Investigation of Pure Coordination Games. Amer. Econ. Rev. 84(3), 658-673.

Nagel, R., 1995. Unraveling in Guessing Games: an Experimental Study. Amer. Econ. Rev. 85(5), 1313-1326

Osborne, M. J., Rubinstein, A., 1998. Games with Procedurally Rational Players. Amer. Econ. Rev. 88(4), 834-847.

Piovesan, M., Wengström, E., 2009. Fast or fair? A study of response times. Econ. Letters 105(2), 193–196.

Rankin, F. W., Van Huyck, J. B., Battalio, R. C., 2000. Strategic Similarity and Emergent Conventions: Evidence from Similar Stag Hunt Games. Games Econ. Behav. 32(2), 315-337.

Rubinstein, A., 1988. Similarity and Decision-Making under Risk (Is There a Utility Theory Resolution to the Allais Paradox?). J. Econ. Theory 46(1), 145-153.

Rubinstein, A., 2007. Instinctive and Cognitive Reasoning: A Study of Response Times. Econ. J. 117(523), 1243-1259.

Rydval, O., Ortmann, A., Ostatnicky, M., 2009. Three Very Simple Games and What It Takes to Solve Them. J. Econ. Behav. Organ. 72(1), 589-601.

Schelling, T. C., 1963. The Strategy of Conflict. New York: Oxford University Press.

Selten, R., Chmura, T., 2008. Stationary Concepts for Experimental 2x2-Games. Amer. Econ. Rev. 98(3), 938-966.

Stahl, D. O., Wilson, P. W., 1994. Experimental Evidence of Players' Models of Other Players. J. Econ. Behav. Organ. 25(3), 309-327.

Stahl, D. O., Wilson, P. W., 1995. On Players' Models of Other Players: Theory and Experimental Evidence. Games Econ. Behav. 10(1), 218-254.

Sugden, R., 1993. Thinking as a Team: Towards an Explanation of Nonselfish Behavior. Social Philosophy and Policy, 10(1), 69-89.

Sugden, R., 1995. A Theory of Focal Points. Econ. J. 105(430), 533-550.

Tversky, A., Kahneman, D., 1981. The Framing of Decisions and the Psychology of Choice. Science 211(4481), 453-458.

Warglien, M., Devetag, G., Legrenzi, P., 1999. I Modelli Mentali dei Giochi: Focalizzazione e Rappresentazioni Selettive. Sistemi Intelligenti XI(1), 85-111.

Weizsäcker, G., 2003. Ignoring the Rationality of Others: Evidence from Experimental Normal-Form Games. Games Econ. Behav. 44(1), 145-171.

Appendices

Appendix A

Equilibrium Analysis

In the article, we used pure strategy Nash equilibria as a benchmark to evaluate observed frequencies. Any manipulation of the descriptive features was always referred to as strategically irrelevant, since it did not erase the starting set of pure strategy Nash equilibria. We now compare the descriptive power of four other stationary concepts, to find which stationary concept best fits our data, and whether any of them can capture effects due to changes in key features.

The stationary concepts tested are: Quantal Response Equilibrium (henceforth QRE; McKelvey and Palfrey, 1995); action sampling equilibrium (Selten and Chmura, 2008); cognitive hierarchy (Camerer et al., 2004); and payoff sampling equilibrium (Osborne and Rubinstein, 1998). Of these, only Nash is non-parametric, whereas the others have one free parameter.

We provide a brief description of parametric stationary concepts. According to QRE (McKelvey and Palfrey, 1995), players make their choices according to relative expected utility and use a quantal choice model. Players also assume that other players apply the same strategy. The possibility of errors in the decision-making process is taken into account.

Action sampling equilibrium is discussed in Selten and Chmura (2008). According to this model, players respond best to a sample (the size of which is the only parameter of the model) of observations of strategies played by their opponents. The parameter is generally set at 7, which is why the model is often considered to be non-parametric. By varying the parameter, we found the value yielding the most accurate fit of our data.

Cognitive hierarchy (Camerer et al., 2004) divides subjects into different strategic categories, according to their level of sophistication. Each subject is assumed to be more sophisticated than the others, and best responds to others‘ behavior by assuming that the other players belong to levels from 0 to k-1 (where k is the level of sophistication of the subject). Types are distributed according to a Poisson distribution.

Payoff sampling (Osborne and Rubinstein, 1998) is similar to action sampling. In this model, players take one sample of actions for each pure strategy available, and then play the strategy with the highest average payoff. This model too has one parameter, since the samples have the same size.

First, we calculated estimates with sample sizes ranging from 1 to 10 for action sampling, and (due to computability restrictions) from 1 to 9 for payoff sampling. We then compared estimated and observed frequencies by the mean square deviation (MSD) and found the parameter value which minimized it. We found optimal sample size parameter values of 9 and 1, for action sampling and payoff sampling, respectively. Similarly, we calculated QRE with values of lambda in the interval 0.01 – 3, with steps of 0.001. For QRE, the parameter value which best fitted the data was 0.096. For QRE estimates, we used a special software: GAMBIT (McKelvey et al., 2010). For the cognitive hierarchy model, the best-fitting parameter was 0.76 (estimate of fitness for values of the parameter ranging from 0 to 10, with steps of 0.01).

Figures A.1a, A.1b, and A.1c show observed and estimated frequencies, divided by row.
In the analysis, together with stationary concepts, we also include the random choice model.

FIGURE A.1a

FIGURE A.1b

FIGURE A.1c

At first sight, Nash and action sampling seem to perform poorly. They generally underestimate the frequency of row 1 (corresponding to strategy HA) and row 2 in matrices with FP. Instead, they overestimate the frequency of row 3, generally corresponding to the equilibrium strategy. Nor do they seem to capture the effects of changes in the variance of HA, whereas Nash cannot capture the effect of FP. Emblematic is the case of DomCol, where both Nash and action sampling give the same estimates in all six versions of the game.
Action sampling often coincides with one of the game Nash Equilibria. When more than one is available, action sampling oscillates between them, and small changes in payoffs can change the expected frequency from 0 to 100%.
Cognitive hierarchy also performs poorly. Although estimates are closer to the observed values, the model does not capture the effects of changes in features, and often maintains estimates invariant in different versions of the same game. In particular, model predictions are not affected in any way by the presence or absence of a focal point.
Payoff sampling clearly performs better than either Nash or action sampling. Even small changes in payoffs affect it, but the reactions are smoother than those observed in action sampling.

Nonetheless, the estimates are not precise, and the difference between estimated and observed frequencies often exceeds 20%.

Of all the stationary concepts, QRE seems to be the best estimator.

Figure A.2 shows MSD scores for stationary concepts and the uniformly distributed random choice model. Since in several games Nash selected more than one prediction, we chose the one closest to the observed frequencies. However, the results show that Nash equilibrium is the worst predictor.

Figure A.2 confirms our observations. There is a clear-cut difference in the accuracy of fit: Nash equilibrium and action sampling equilibrium perform poorly, whereas cognitive hierarchy, payoff sampling and QRE perform significantly better. Random choice falls between the two groups, outperforming Nash and action sampling. However, the trend of the data shown in Figures A.1a, A.1b, and A.1c indicates that the first is probably the result of a statistical artifact.

Differences in performances were tested by a two-tailed t-test[15]. We compared the observed frequencies for each matrix row with the estimates of the stationary concepts and of the uniformly distributed random choice model. The statistical analysis confirms our previous results: QRE performs significantly better than Nash, random choice, action sampling, cognitive hierarchy ($p = 0$), and payoff sampling ($p \leq 0.1$). The second-best model is payoff sampling, which performs better than Nash and action sampling ($p = 0$), and random choice ($p = 0.01$), but not cognitive hierarchy. Cognitive hierarchy performs significantly better than Nash ($p = 0$), action sampling ($p = 0.01$), and random ($p \leq 0.1$). Random choice performs better only than Nash ($p \leq 0.05$), whereas Nash and action sampling are statistically indistinguishable.

FIGURE A.2

Concluding, as suggested by the analysis of aggregate choices, Nash equilibrium performs poorly and captures almost none of the effects of the descriptive features. Of all the other stationary concepts analyzed, QRE is the best estimator. This result is quite interesting, as in previous studies (e.g., Selten and Chmura, 2008) QRE was the second-worst performer, better only than Nash. With the features we take into consideration, QRE is able to capture even minute modifications, avoiding overreactions.

---

[15] Similar results were obtained with a two-tailed Wilcoxon signed rank test.

Appendix B

Instructions for the experiment

INSTRUCTIONS

Welcome!

You are about to participate in an experiment on interactive decision-making, funded by the R.O.C.K. (Research on Organizations, Coordination and Knowledge) research group of the University of Trento. Your privacy is guaranteed: results will be used and published anonymously.

All your earnings during the experiment will be expressed in **Experimental Currency Units** (ECUs). Your earnings will depend on your performance in the experiment, according to the rules which we will explain to you shortly. You will be paid privately and in cash at the end of the experimental session. Other participants will not be informed about your earnings.

The experiment is divided in two, unrelated parts. The instructions for the second part will be distributed at the end of the first part. Your behavior and the earnings you obtain in the first part do not affect your earning in the second part in any way. The maximum you can earn in the experiment is 20 Euros.

PART 1

The experiment consists of 30 rounds; in each round you will face an interactive decision-making situation. The word “interactive” means that the outcome of your decision will be determined by your choice and by the choice of another participant, randomly chosen. More specifically, your earnings in each decision-making situation will be determined by the combination of your choice and the choice of the participant with whom you will be paired in that round.

EXPERIMENTAL STRUCTURE

The structure of each interactive decision problem, henceforth GAME, will be represented by a table like the one below:

		OTHER PLAYER’S ACTIONS (Column Player)	
		C1	C2
YOUR ACTIONS (Row Player)	R1	(6,4)	(4,7)
	R2	(3,4)	(5,6)

The table is to be read as follows: you and the participant with whom you are paired will play the roles, respectively, of ROW PLAYER and COLUMN PLAYER, or vice versa. The available choices of the ROW PLAYER are represented by the rows of the table (in the example, R1 and R2),and the available choices of the COLUMN PLAYER are represented by the columns of the table (in the example, C1 and C2).

If your role in a round is that of ROW PLAYER, the participant with whom you are paired will have the complementary role of COLUMN Player, and vice versa. You will learn your role by reading the labels on the table. The label “YOUR ACTIONS” will be placed close to your role, and the label “OTHER PLAYER’S ACTIONS” will be close to the role of the player you are paired with. For example, in a table like the one presented above, you have the role of ROW player, and the player with whom you are paired has the role of COLUMN player, so that the labels are inverted.

IMPORTANT: you will keep the same role (ROW or COLUMN) in all the decisional tables of the experiment, although the participant with whom you are paired will be picked randomly (and therefore may be different) in each round.

Each possible combination of choices of row and column player (i.e., each possible combination of rows and columns of the table) identifies one cell in the matrix. Each cell reports two numerical values in brackets. These values indicate the earnings (in Experimental Currency Units) of each participant associated with that combination of choices. Conventionally, the first number represents the earnings of the ROW PLAYER (regardless of whether it is you or the other player), and the second number represents the earnings of the COLUMN PLAYER.

For example: in the table below, if YOU, the ROW PLAYER, choose row R1 and the OTHER PLAYER chooses column C2, then your earnings will be those in the cell at the intersection between row R1 and column C2; YOU (ROW Player) earn 4 ECUs and the OTHER PLAYER (COLUMN PLAYER) 7 ECUs.

		OTHER PLAYER (Column Player)	
		C1	C2
YOU (Row Player)	R1	(6,4)	(4,7)
	R2	(3,4)	(5,6)

Bear in mind that you cannot directly choose the cell of the table, but only one of the rows or columns, depending on your role. Only the combination of both choices will select one and only one cell, corresponding to your earnings and to those of the other participant.

MATCHING RULES

For each decisional table, the participant with whom you are paired is randomly selected by the software. Obviously, as the matching rule is random and as the number of decisional tables larger than the number of participants in the session, during the experiment you will be paired more than once with the same subject. However, you will never know the identity of the participant you are matched with, nor will you know that person's choice in a table after you have made yours.

INFORMATION

In each of the 30 rounds, the screen will show the decisional table for that round, and you will be asked to make a decision. Each table is marked by a numerical code, which will be used for the final payment. The code appears in the top left-hand corner of each decisional table. The top right-hand corner of the screen specifies the time remaining for your decision. You must communicate your decision by typing 1, 2 or 3 in the space “I choose row/column number”, and by clicking the “confirm” button with the mouse.

In order for the next round to start, ALL participants must have entered their decision for the current round, and we therefore ask you not to take more than 30 seconds to choose. After 30 seconds, a text message in the top right-hand corner of the screen will ask you to write down your decision. If you delay your decision considerably, you will oblige the other players to wait.

You will face 30 decisional matrices, corresponding to 30 different interactive situations. There is no relation among your choices in the different games, each game is independent of the others. At the end of the 30th round, the first part of the experiment will be completed, and your earnings for this part will be determined.

PAYMENTS

Each matrix is identified by a code. Some tags have been placed in a box, each showing the code of one of the matrices. The experimenter will ask one of you, selected randomly, to verify that the box contains 30 tags, and also that the codes on the tags are really different from each other. Subsequently, the experimenter will ask a different participant, selected randomly, to pick 5 of these tags from the box. Each of you will be paid according to the earnings obtained in the tables corresponding to the extracted codes. The earnings in each of the 5 selected tables will be determined by matching your choice with the choice of the participant with whom you were matched at that table. Since each of the 30 decisional tables of the experiment has a positive probability of being selected for payment, we ask you to devote the same attention to all of them.

Before the experiment starts, we will ask you to answer a simple anonymous questionnaire, in order to make sure that you have understood the instructions perfectly or whether clarifications are needed. If there are incorrect answers, the relevant part of the instructions will be repeated. After the questionnaire phase is completed, the experiment will start.

It is very important that you remain silent during the experiment, and that you never communicate with the other participants, either verbally, or in any other way. For any doubts or problems you may have, please just raise your hand and the experimenter will approach you. If you do not remain silent or if you behave in any way that could potentially disturb the experiment, you will be asked to leave the laboratory, and you will not be paid.

Thank you for your kind participation!

Appendix C

QUESTIONNAIRE

Dear Participant,

The following questionnaire is anonymous and has the sole purpose of verifying your understanding of the rules of this experiment.

We ask you to answer to the following questions. If you are uncertain about how to respond, please consult the instructions sheet.

When you have finished, please raise your hand and a member of the staff will check that all your answers are filled in.

Thank you for your cooperation!

COLUMN Player

ROW Player

	C1	**C2**	**C3**
R1	10,20	30,40	50,40
R2	1,2	3,4	6,3
R3	15,30	5,9	15,7

Suppose you are assigned the role of ROW PLAYER: If the COLUMN PLAYER chooses strategy C2 and you choose strategy R2, how many ECUs will you earn? .... And the other player?...........If you choose strategy R2, and COLUMN PLAYER chooses strategy C3, how many ECUs will that person earn? .... And what about you? ....

If the other player chooses C1, your earnings will be:

If you choose R1: ....

If you choose R2: ....

If you choose R3: ....

Suppose you are assigned the role of COLUMN PLAYER

If the ROW PLAYER chooses strategy R2 and you choose strategy C1, how many experimental points will you earn? .... And the other player?...........

If the other player chooses R1, your earnings will be:

If you choose C1: ....

If you choose C2: ....

If you choose C3: ....

Your role (as ROW or COLUMN PLAYER) in the rounds of the experiment will change:

TRUE or FALSE

The participant with whom you are paired will be determined randomly in each round, and you will never be matched more than once with the same participant.

TRUE or FALSE

After you have taken your decision on a table, you will be able to observe the choice of the participant with whom you were paired.

TRUE or FALSE

Appendix D

Instructions for Experiment (Phase 2)

The sheet given to you shows 10 numbered ROWS, and each ROW presents 2 OPTIONS: **L** and **R**. We ask you to choose one and only one of the two options in each row. Your earnings will be determined in the following way.

This is a box containing 10 numbers, from 1 to 10, which will be used to determine your earnings. After you have made your choices, we will extract 2 numbers: the first number will determine the ROW that will be used to calculate your earnings, and the second number will determine your earnings given the OPTION, L or R, that you chose for that ROW. Obviously, each ROW has the same probability of being chosen, i.e., 1 of out 10.

Now, pay attention to ROW 1. OPTION L pays 2 Euros if the number drawn is 1, and 1.60 Euros if the number drawn is a number between 2 and 10 (extremes included). OPTION R pays 3.85 Euros if the number drawn is 1, and 0.1 Euros if the number drawn is a number between 2 and 10 (extremes included). All the ROWS are similar, meaning that the earnings for both OPTIONS remain the same. The only difference is that, moving towards the bottom of the table, the possibility of winning the larger amount increases for both OPTIONS. Consequently, the possibility of winning the lower amount decreases. If ROW 10 is selected, there will be no need to extract the second number, because each OPTION will certainly pay the larger amount, that is, 2 Euro (et seq.) for OPTION L and 3.85 Euros for OPTION R.

L is the default option for all ROWS, but you can choose to switch to OPTION R by simply marking the desired ROW. If you prefer OPTION R from a certain point onwards, just mark the corresponding ROW. Please note that you can switch from L to R only once and that the switch is irreversible; therefore, you must mark only ONE ROW, which indicates that, in all the ROWS above, you prefer OPTION L, whereas in the marked ROW and in all ROWS below, you prefer OPTION R. If you do not want to change, i.e., if you prefer OPTION L in all ROWS, don't mark anything. If you always prefer OPTION R, you must mark the first ROW. You can choose any of the 10 ROWS, but you can only pass from L to R once, and therefore at most you can put 1 mark.

When you have finished, we will collect your sheet. When all participants have completed their choices, one of you will draw the two numbers from the box. Remember, the first extraction determines the ROW that will be used to calculate everybody's earnings, and the second number will determine your earnings; the first number will be put back in the box before the second number is extracted. Your earnings in this choice task will be added to those obtained in the first part of the experiment, and the total amount will be paid to you privately at the end of the experiment.

EXAMPLE

Suppose that the ROW drawn randomly is ROW 3, and that you have marked one of the rows below ROW 3. Since ROW 3 is above your mark, this indicates that you prefer OPTION L for ROW 3. Then, if the second drawn number is (for example) 5, your earnings are 1.6 Euros.

Please answer the questions at the end of the sheet. We need this information for statistical purposes only.

	Option L	**Switch from L to R**	**Option R**
ROW 1	2 € with 1 or 1.6 € with 2-10		3.85 € with 1 or 0.1 € with 2-10
ROW 2	2 € with 1-2 or 1.6 € with 3-10		3.85 € with 1-2 or 0.1 € with 3-10
ROW 3	2 € with 1-3 or 1.6 € with 4-10		3.85 € with 1-3 or 0.1 € with 4-10
ROW 4	2 € with 1-4 or 1.6 € with 5-10		3.85 € with 1-4 or 0.1 € with 5-10
ROW 5	2 € with 1-5 or 1.6 € with 6-10		3.85 € with 1-5 or 0.1 € with 6-10
ROW 6	2 € with 1-6 or 1.6 € with 7-10		3.85 with 1-6 or 0.1 € with 7-10
ROW 7	2 € with 1-7 or 1.6 € with 8-10		3.85 € with 1-7 or 0.1 € with 8-10
ROW 8	2 € with 1-8 or 1.6 € with 9-10		3.85 € with 1-8 or 0.1 € with 9-10
ROW 9	2 € with 1-9 or 1.6 € with 10		3.85 € with 1-9 or 0.1 € with 10
ROW 10	2 € with 1-10		3.85 € with 1-10

Please answer the following questions:

What faculty are you enrolled in? ____
When did you enrol? (year)
When were you born? _________/__________/________

Please specify where you were born and your nationality __________________________________

____

Specify M or F

Have you attended any courses on Game Theory?

____

If so, which courses? ____

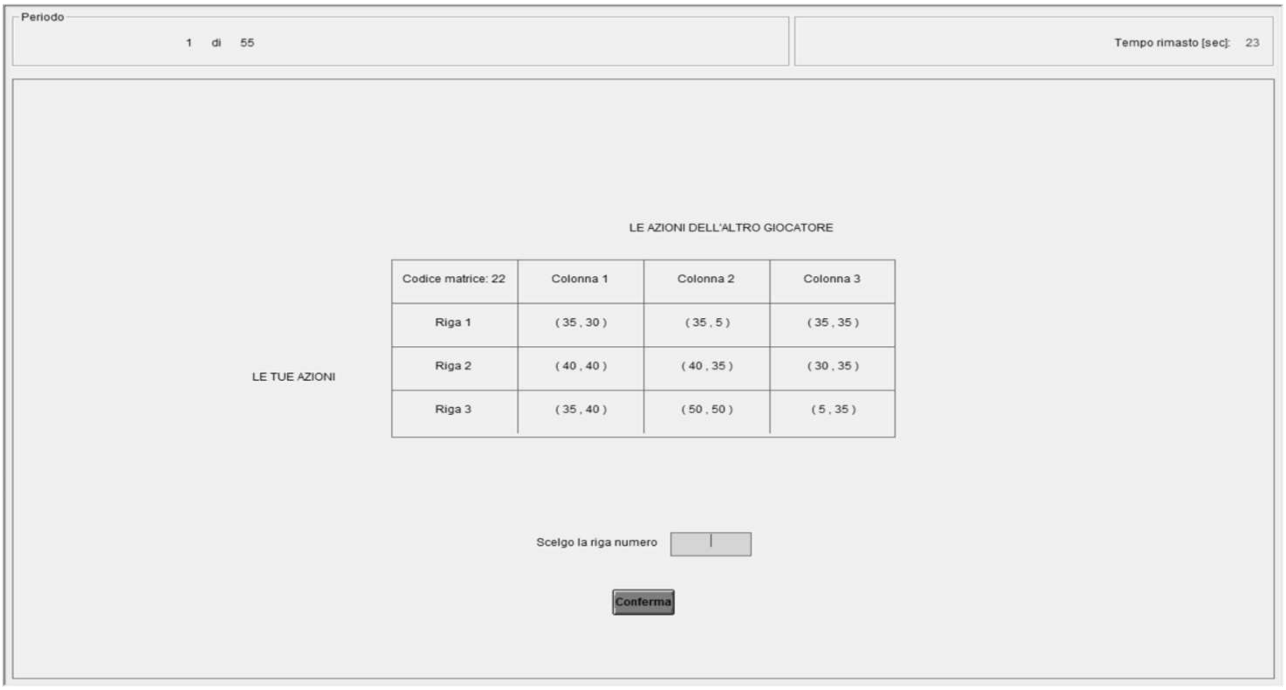
Do you know what a Nash Equilibrium is? ____

If so, in what courses did you study it?

____

Figures

Figure 1



Periodo

1 di 55

Tempo rimasto [sec]: 23

LE AZIONI DELL'ALTRO GIOCATORE

LE TUE AZIONI

Codice matrice: 22	Colonna 1	Colonna 2	Colonna 3
Riga 1	(35 , 30)	(35 , 5)	(35 , 35)
Riga 2	(40 , 40)	(40 , 35)	(30 , 35)
Riga 3	(35 , 40)	(50 , 50)	(5 , 35)

Scelgo la riga numero

Fig. 1: Example of the software interface as it was presented to subjects in the instructions

Figure 2a

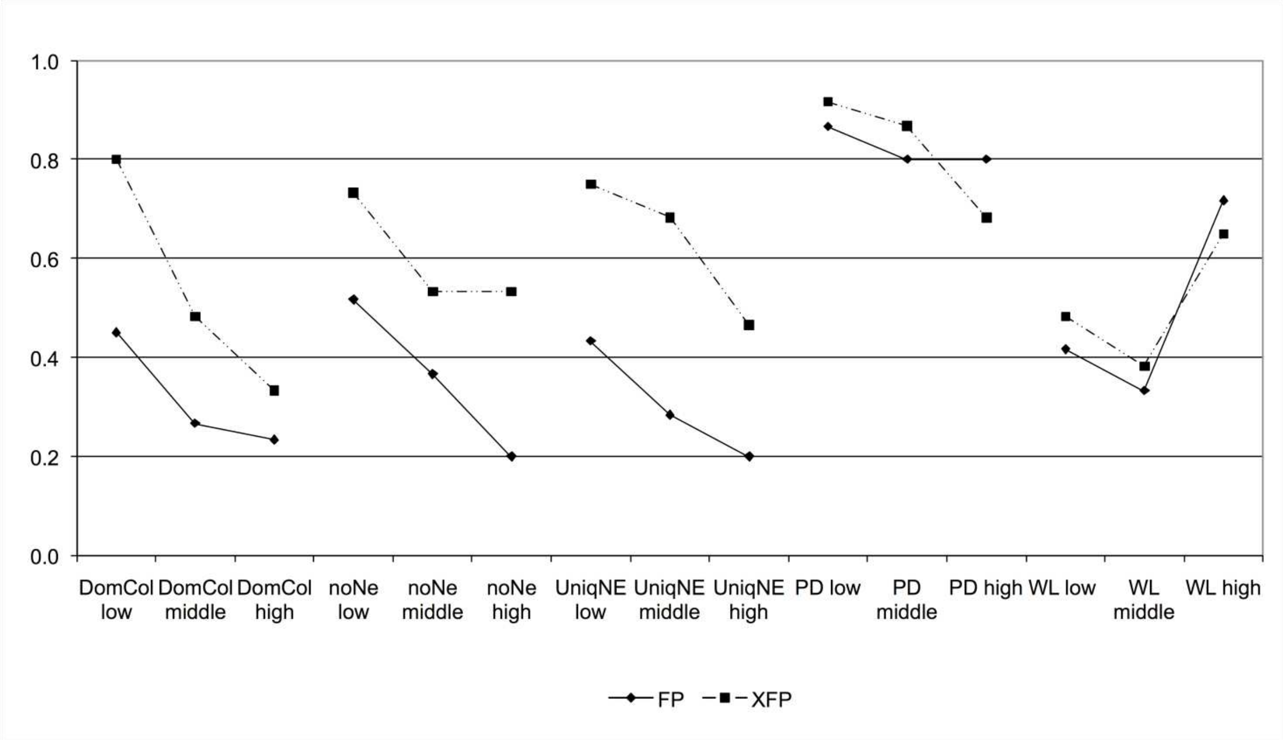


Fig. 2a: Observed frequencies of row 1 choices

Figure 2b

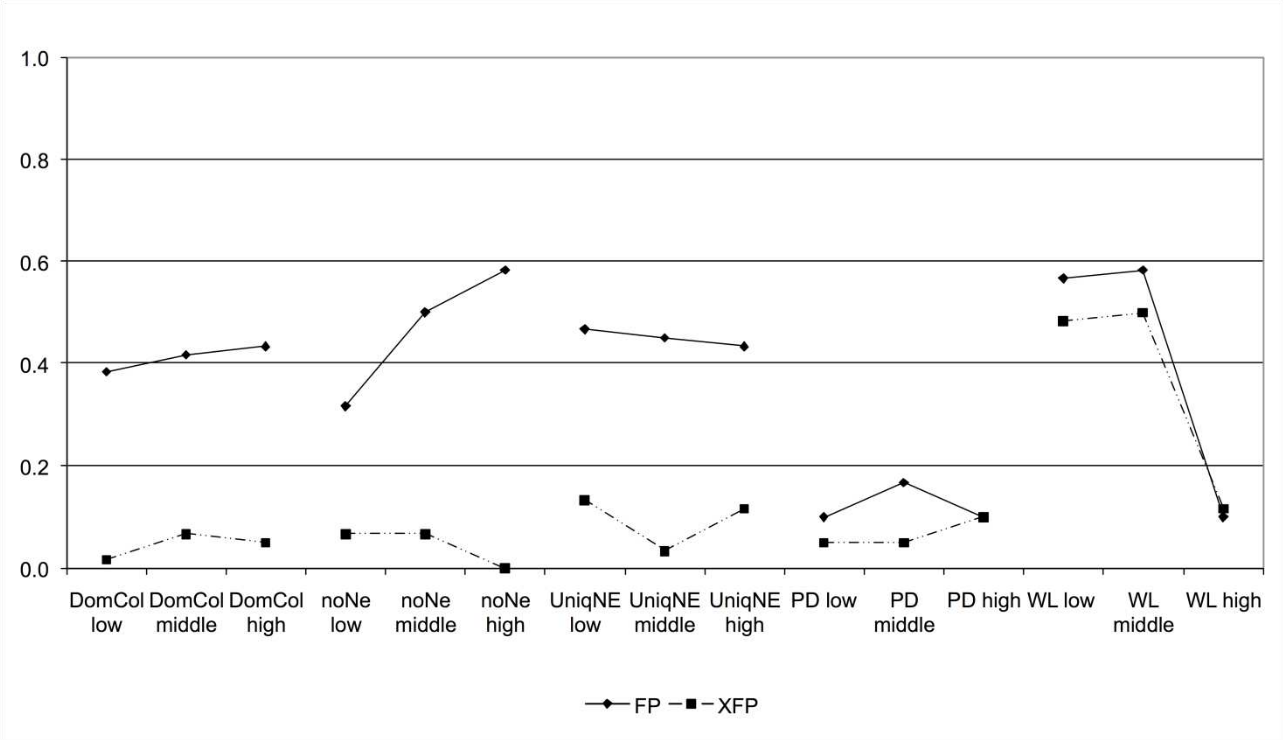


Fig. 2b: Observed frequencies of row 2 choices

Figure 2c

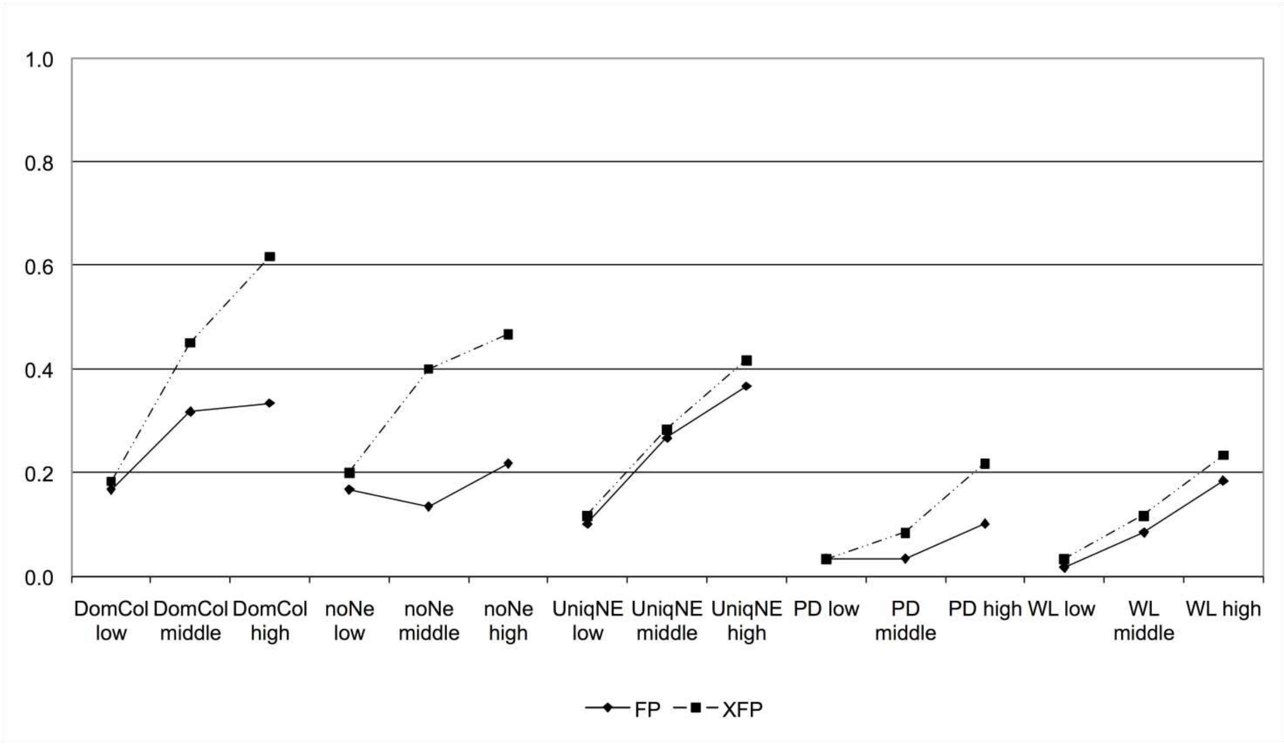


Fig. 2c: Observed frequencies of row 3 choices

Figure 3

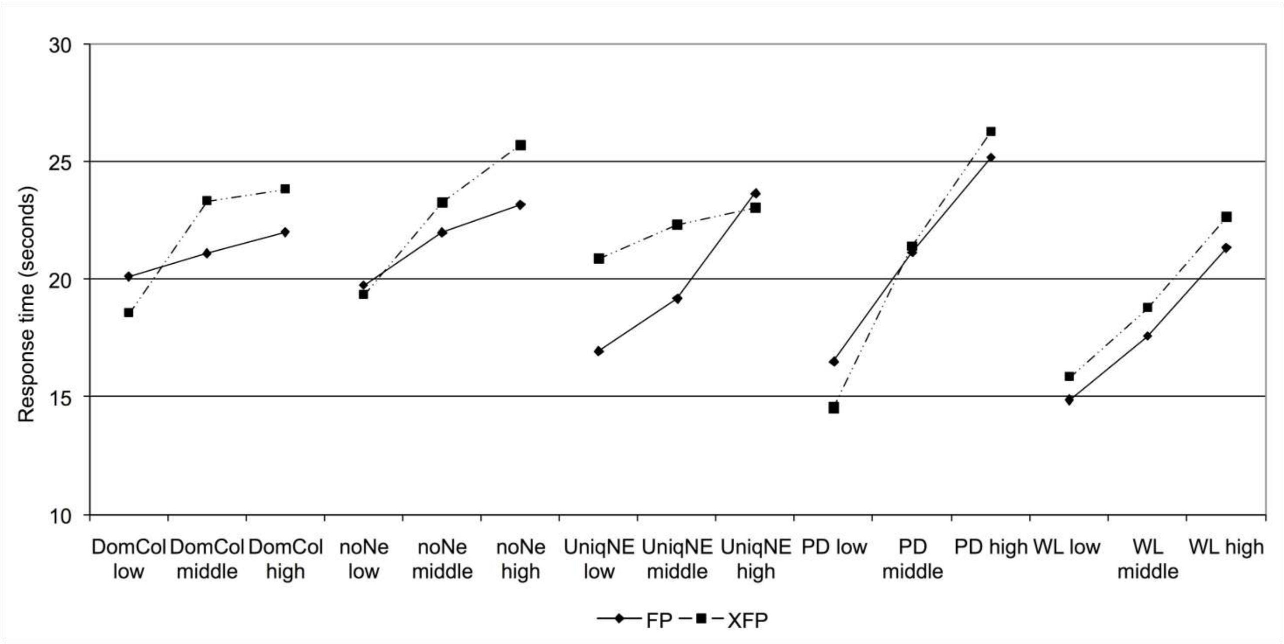


Fig. 3: Average response time in seconds, for each matrix

Figure 4

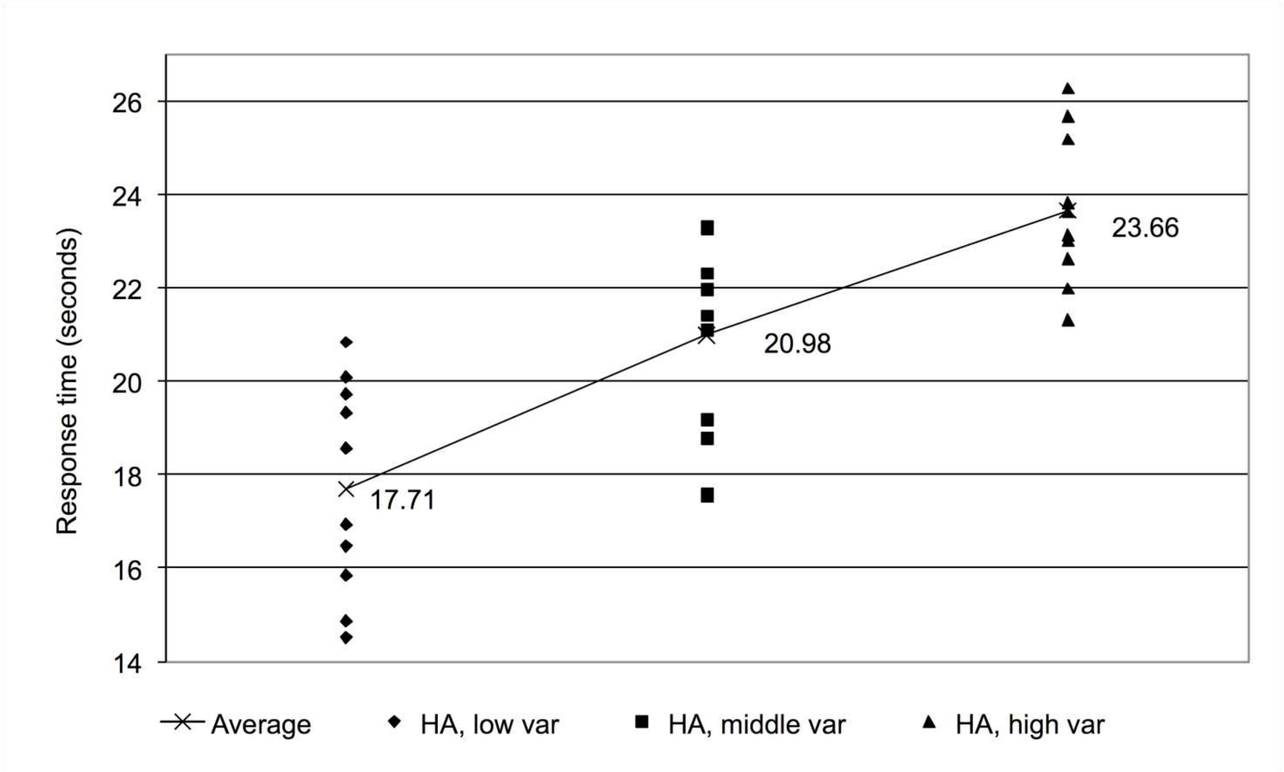


Fig. 4: Average response time as a function of HA variance level

Figure A.1a

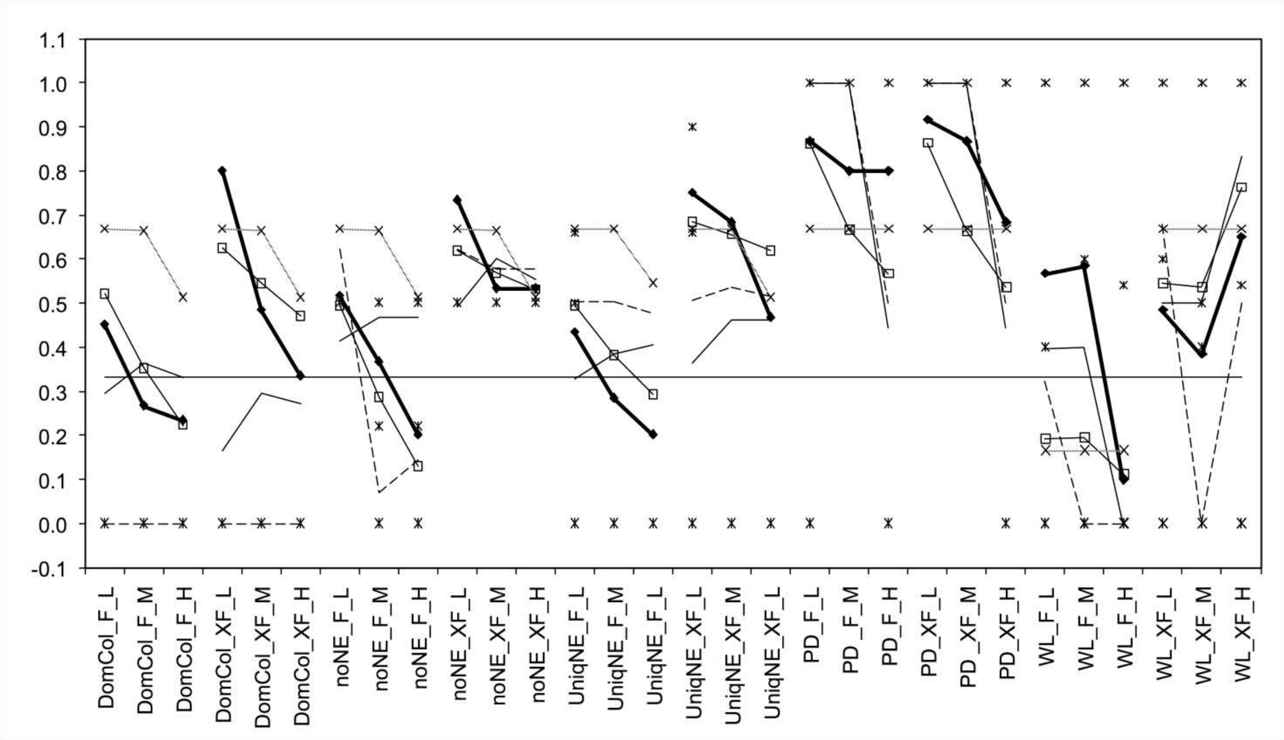


Fig. A.1a: Observed and estimated frequencies for row 1 choices: Nash Equilibria (stars), Action Sampling (dashed line), Payoff Sampling (thin continuous line), QRE (thin continuous line, with empty squares), Cognitive Hierarchy (dotted line), Random Choice (continuous horizontal line), Observed Frequencies (thick continuous line, with small squares)

Figure A.1b

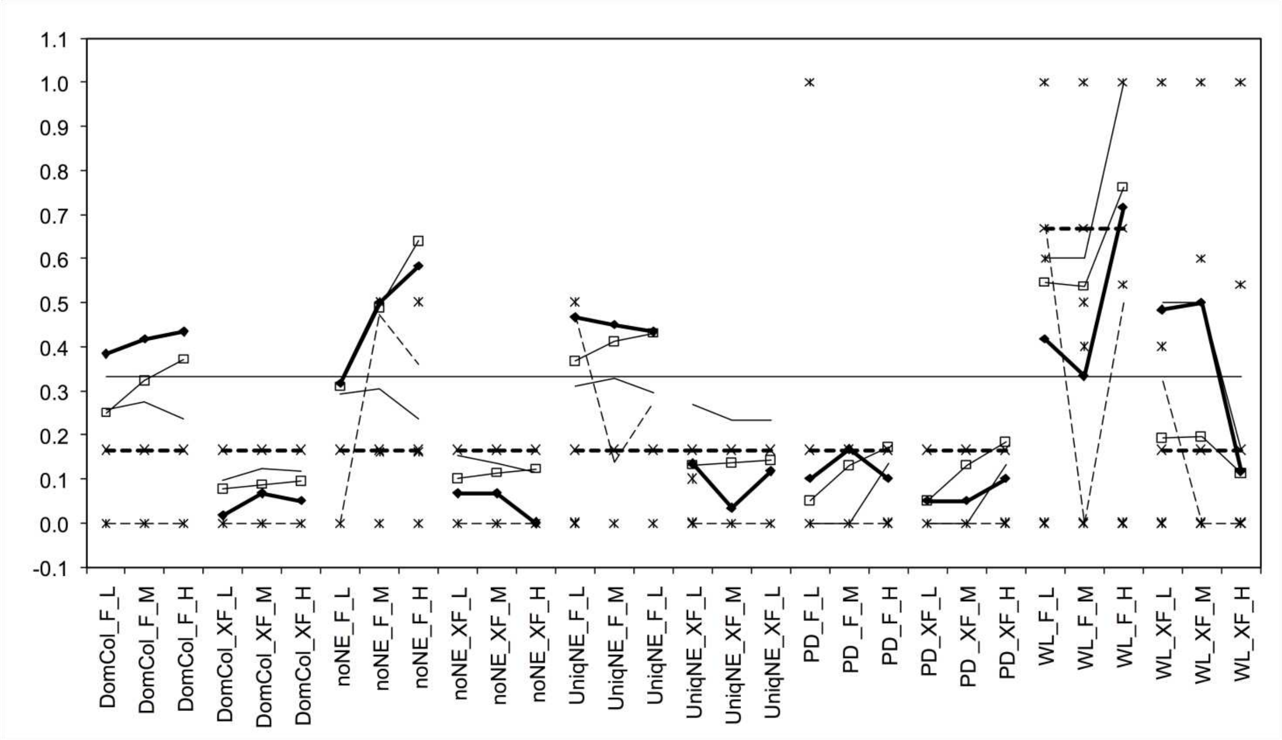


Fig. A.1b: Observed and estimated frequencies for row 2 choices.

Figure A.1c

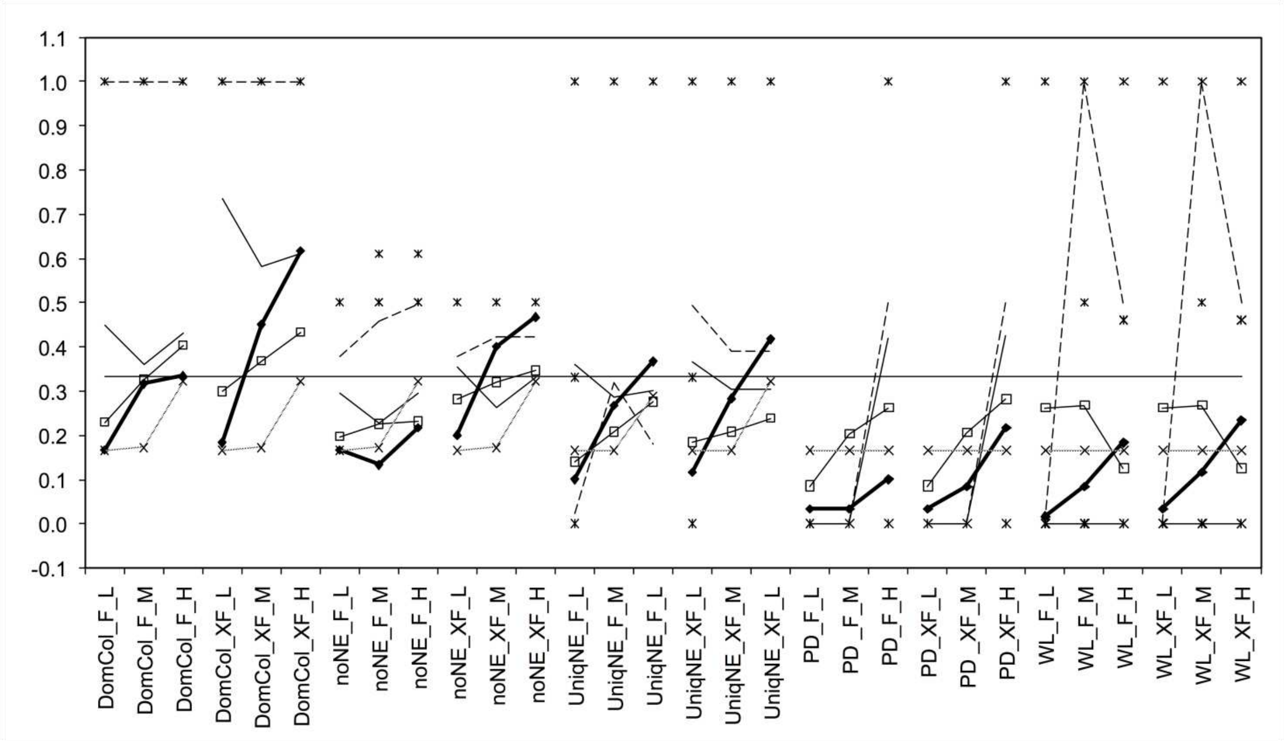


Fig. A.1c: Observed and estimated frequencies for row 3 choices.

Figure A.2

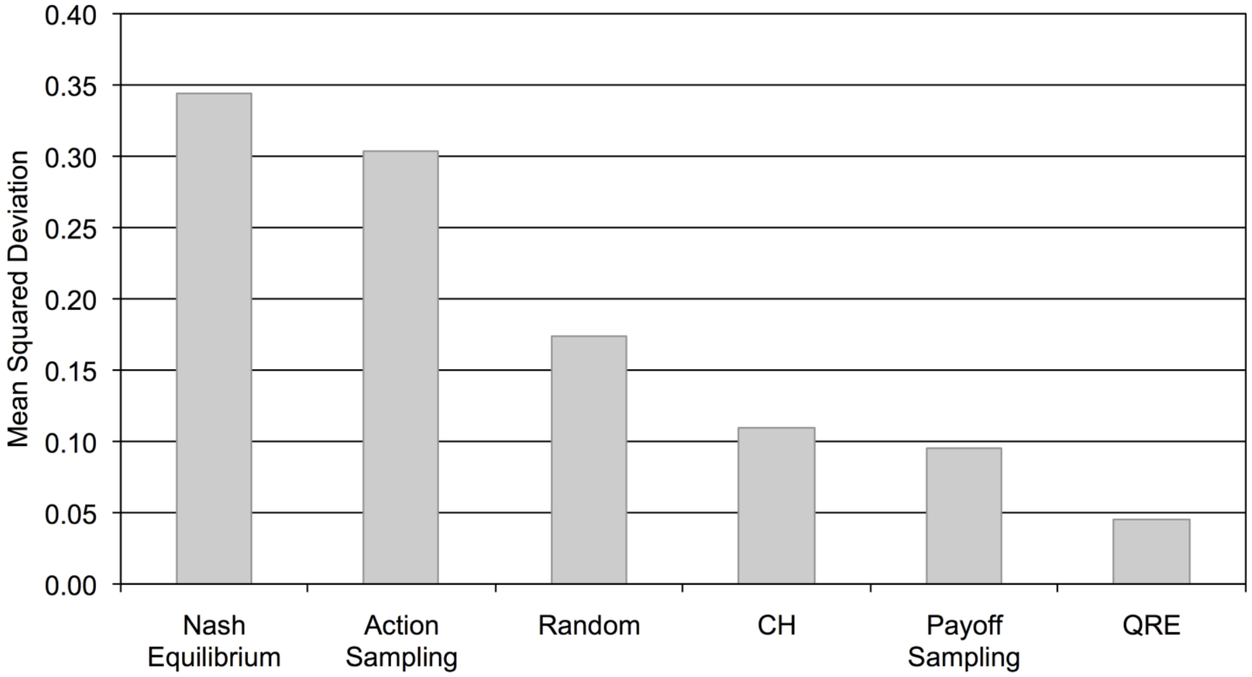


Fig. A.2: Overall mean squared distances of five stationary concepts