HAL
open science

Cognitive ability and the effect of strategic uncertainty

Nobuyuki Hanaki, Nicolas Jacquemet, Stéphane Luchini, Adam Zylbersztejn

To cite this version:

Nobuyuki Hanaki, Nicolas Jacquemet, Stéphane Luchini, Adam Zylbersztejn. Cognitive ability and the effect of strategic uncertainty. 2014. halshs-01095897

HAL Id: halshs-01095897

https://shs.hal.science/halshs-01095897v1

Preprint submitted on 16 Dec 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

amse
aix marseille
school of economics

Working Papers / Documents de travail

Cognitive Ability and the Effect of Strategic Uncertainty

Nobuyuki Hanaki
Nicolas Jacquemet
Stéphane Luchini
Adam Zylbersztejn

WP 2014 - Nr 58

Aix*Marseille université
cnrs
L'ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES
Inserm
IRD Institut de recherche pour le développement
Centrale Marseille

Cognitive ability and the effect of strategic uncertainty*

Nobuyuki Hanaki† Nicolas Jacquemet‡ Stéphane Luchini§ Adam Zylbersztejn¶

December 16, 2014

Abstract

How is one's cognitive ability related to the way one responds to strategic uncertainty? We address this question by conducting a set of experiments in simple 2×2 dominance solvable coordination games. Our experiments involve two main treatments: one in which two human subjects interact, and another in which one human subject interacts with a computer program whose behavior is known. By making the behavior of the computer perfectly predictable, the latter treatment eliminates strategic uncertainty. We find that subjects with higher cognitive abilities are more sensitive to strategic uncertainty than those with lower cognitive abilities.

Keywords: Strategic Uncertainty, Robot, Bounded Rationality, Experiment

JEL Classification: C92, D83.

*This project is partly financed by JSPS-ANR bilateral research grant "BECOA" (ANR-11-FRJA-0002). Ivan Ouss provided much efficient research assistance. We thank Juergen Bracht, Colin Camerer, Guillaume Fréchette, Asen Ivanov, Frédéric Koessler, Rosemarie Nagel, Ariel Rubinstein, Jason F. Shogren, Jean-Marc Tallon, Antoine Terracol and Marie-Claire Villeval for their comments. Hanaki and Jacquemet gratefully acknowledge the *Institut Universitaire de France*. Luchini thanks the School of Business at the University of Western Australia for hospitality and support.

†Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS. 2 Rue de la Charité. 13002, Marseille, France. nobuyuki.hanaki@univ-amu.fr

‡Université de Lorraine (BETA) and Paris School of Economics. 3 Place Carnot, 54035 Nancy. Nicolas.Jacquemet@univ-lorraine.fr

§Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS. 2 Rue de la Charité. 13002, Marseille, France. stephane.luchini@univ-amu.fr

¶Université de Lyon, Lyon, F-69007, France; CNRS, GATE Lyon Saint-Etienne, 93 Chemin des Mouilles, F-69130 Ecully, France. zylbersztejn@gmail.com.

1 Introduction

Coordination games provide a useful game-theoretical paradigm for analyzing a wide range of economic phenomena, such as macroeconomic fluctuations (Cooper and John, 1988), bank runs and speculative currency attacks on financial markets (Morris and Shin, 2003; Heinemann, 2012), and commercial production processes (Brandts, Cooper, and Weber, 2014). Because of the multiplicity and the Pareto-rankability of the Nash equilibria in these games, and because decisions are usually made in a state of strategic uncertainty regarding others' behavior, the resulting outcomes can be driven away from the Pareto-Nash equilibrium – a phenomenon known as coordination failure.

Coordination failure has been shown to be a persistent pattern in numerous lab implementations (Camerer, 2003, Ch.7). The present paper contributes to a large body of experimental studies exploring this welfare-reducing phenomenon. A vast part of this literature is based on the core idea that coordination failures arise from strategic uncertainty, and various institutional designs are put forth as a remedy against it: introducing repeated encounters, varying the stability and the size of groups, providing information feedback, allowing for observation of others' past behavior, or introducing pre-play communication between players (see also Devetag and Ortmann, 2007, for an extensive survey of this literature.) However, although these mechanisms are usually found to improve efficient coordination, they fall short of completely solving the problem of coordination failure.

In this paper, we take a further step to deepen our understanding of the nature of coordination failure. Our experimental results confirm that strategic uncertainty is an important determinant of the efficiency of strategic decision-making. Even more importantly, our experiment shows that individual cognitive abilities have a strong link with the way strategic uncertainty influences the decisions of subjects in our experimental coordination games.

Our investigation involves a classic 2×2 coordination game, based on Selten (1975) and Rosenthal (1981), and presented in Table 1. With $L < S < H$, $m < h$, and $s < h$, the game is one-step dominance solvable: the elimination of Player B's weakly dominated strategy l immediately leads to the Pareto-Nash equilibrium (R, r). Moreover, from the standard theory perspective, (R, r) is a natural candidate for a focal point, since it is also risk-dominant.[1]

Notwithstanding these predictions, various studies have found a frequent failure to achieve the efficient equilibrium (see, e.g., Beard and Beil, 1994; Beard, Beil, and Mataga, 2001; Goeree and Holt, 2001; Cooper and Van Huyck, 2003; Jacquemet and Zylbersztejn, 2014) both in sequential and simultaneous implementations of this game. Depending on the exact experimental setup, between 20% and 84% of observed outcomes are not Pareto efficient. While the literature has long focused on the strategic uncertainty faced by Player A as the source of coordination failures

[1] Another Nash equilibrium, (L, l), involves a weakly dominant strategy l by Player B. The existence of a clear-cut theoretical benchmark distinguishes this game from another well-known 2x2 coordination game, the stag hunt, in which each Nash equilibrium is supported by a certain solution concept – either payoff-dominance or risk-dominance.

Table 1: A simple two-person two-action dominance solvable coordination game

		Player B	
		l	r
Player A	L	(S ; s)	(S ; s)
	R	(L ; m)	(H ; h)

in these experiments, recent evidence provides different clues for explaining this behavior. For example, Polonio, Di Guida, and Coricelli (2014) use eye-tracking data gathered from simple 2×2 games to demonstrate that some subjects do not pay attention to the payoffs of their opponent, and thus do not realize that the opponent has a dominant strategy. Thus, some Player As in our coordination game may choose L without taking Player Bs' behavior into consideration at all – which precludes any meaningful role of strategic uncertainty on the decision-making of the former.

Therefore, the first aim of this paper is to understand the extent to which deviations from strategy R by Player As is due to strategic uncertainty, which constitutes an important step towards designing more efficient mechanisms aimed at eliminating coordination failure. To address this issue we conduct a set of experiments based on four variations of a dominance solvable coordination game shown in Table 1, in which human subjects (acting as Player As) interact with Player Bs represented by either (a) other human subjects, or (b) a computer program. Computerized Player Bs are programmed to always choose r, and this fact is clearly explained to the subjects. Therefore, subjects acting as Player As interacting with computers do not face any strategic uncertainty, which provides an empirical benchmark for assessing the effect of strategic uncertainty on Player As' behavior in human-human interactions.

In this sense, our experiment is related to a recent and growing body of experimental studies that seek to separate and evaluate the behavioral effect of strategic uncertainty in collective decision-making. For example, in an alternating bargaining game, Johnson, Camerer, Sen, and Rymon (2002) investigated the effects of two potential causes for failure in backward induction: confusion or other-regarding social preferences. They found evidence that confusion was an important cause of deviations from the equilibrium outcome. Houser and Kurzban (2002) and Ferraro and Vossler (2010) do the same in public good contribution experiments, and estimate that confusion explained up to around one half of contribution levels. Fehr and Tyran (2001) focused on the strategic aspects of "nominal illusion". They considered four-player repeated price setting games, and introduced a negative nominal shock in the middle of the experiment. They found that roughly half of non-immediate adjustment to the new equilibrium after the shock was due to individual bounded rationality (or confusion) and the other half was due to strategic uncertainty.

Finally, Akiyama, Hanaki, and Ishikawa (2013) investigated the magnitude of the effect of strategic uncertainty in explaining the observed deviation of price forecasts from the fundamental values in an experimental asset market *à la* Smith, Suchanek, and Williams (1988). They found significant effects of both confusion and strategic uncertainty.[2]

Our second objective is to shed new light on the relationship between cognitive ability and strategic thinking. In particular, we investigate whether the failure to seek efficiency by choosing R is more widespread among Player As with low cognitive abilities than for those with high cognitive abilities. We address this question by conducting a cognitive ability test in several experimental sessions (involving both human-human and human-robot interactions), and correlate the efficiency of observed behavior and subjects' test scores, while controlling for the presence of strategic uncertainty.

From this perspective, our study contributes to recent literature that investigates the relationship between subjects' cognitive abilities and their degree of strategic sophistication. For example, Brañas-Garza, García-Muñoz, and Hernán (2012) reported that subjects with higher scores on the Cognitive Reflection Test (CRT, Frederick, 2005) choose, on average, numbers closer to the Nash equilibrium in the beauty contest games. In the same vein, Akiyama, Hanaki, and Ishikawa (2013) reported that the magnitude of the effect of strategic uncertainty is positively correlated with subjects' scores on the CRT test, while the effect of confusion is negatively correlated with the score. Burks, Carpenter, Goette, and Rustichini (2009) reported that subjects (trainee truckers) with higher scores in Raven's progressive matrix test[3] are more patient and more willing to take calculated risks. In addition, they report that subjects with higher Raven's test scores more accurately predict others' behavior in a sequential prisoners' dilemma game, and better adapt their behavior to others' behavior. Carpenter, Graham, and Wolf (2013) showed that subjects with higher scores in Raven's test more frequently win in "Race to 5, 10, or 15" games[4] and guessed others' choices better in a 20-player beauty contest game. Finally, Gill and Prowse (2013) also reported that subjects with higher scores in Raven's test not only choose numbers closer to the equilibrium in a repeated 3-player beauty contest game, but also respond to the average score of other subjects in the group by choosing number close to the equilibrium when facing with others with higher scores

[2]Other studies, somewhat less related to ours, used robots that do not follow equilibrium strategies as a way to control for subjects' beliefs about the behavior of their opponents. For instance, Ivanov, Levin, and Niederle (2010) used robots to replicate the observed behaviors of subjects in the past, and Embrey, Fréchette, and Lehrer (2014) and Costa-Gomes and Crawford (2006) used robots to make some players follow the predetermined distribution of boundedly rational behaviors.

[3]Raven's test is widely used to measure "fluid intelligence", that is "the capacity to think logically, analyse and solve novel problems, independent of background knowledge" (Mullainathan and Shafir, 2013, p.48).

[4]The "Race to 5 (or 10 or 15)" game is a two player sequential move game in which two players, moving alternatively, can put either 1, 2 or 3 stones in a common hat which is empty at the beginning. The player who puts the 5th (or 10th or 15th, respectively) stone in the hat wins. The first mover has a clear advantage in this game, and one can derive the winning strategy by a backward induction. The difficulty of deriving the winning strategy increases with the number of target stones.

than when facing with others with lower scores. To sum up, these empirical studies suggest that people with high cognitive abilities respond more aptly to strategic conditions they face than those with low cognitive abilities. The present study extends this investigation to a new and important economic environment – the coordination game.

As will be seen, we find that Player As' failure to choose R can be only partially explained by uncertainty about their partners' intentions: in many cases, the former act in this manner even when interacting with a computer program that is known to always act reliably by choosing r. We also report that Player As with high cognitive abilities (measured in terms of Raven's test scores) tend to be more sensitive to strategic uncertainty than those with low cognitive abilities.

2 Experimental design

We consider four payoff matrices based on the simultaneous-move coordination game shown in Table 1. Our main manipulation lies in varying the nature of Player B, who may be represented either by a human subject (Human treatment) or a pre-programmed computer (Robot treatment). All games and treatments are implemented using a between-subject design – only one version of the game is played in each experimental session. In all the sessions, the one-shot game is repeated ten times with participants' roles remaining fixed, pairs being rematched in each round using a perfect stranger, round-robin procedure[5] and take-home earnings corresponding to a single round randomly drawn at the end of each experimental session. At the end of each round, subjects are only informed of their own payoffs.

2.1 Treatments and hypotheses

The four experimental game matrices are presented in Table 2. Two of them have already been experimentally studied in the literature. Baseline Game 1 (shown on the top left panel of Table 2) was used as the baseline treatment in Beard and Beil (1994); Beard, Beil, and Mataga (2001); Jacquemet and Zylbersztejn (2013, 2014). Egalitarian Game 2 (shown on the bottom right panel of Table 2) was one of the additional matrices introduced by Jacquemet and Zylbersztejn (2014) in an attempt to assess the effect of the relative payoff structure on subjects' behavior. The latter study reports a strong divergence between both players' behavior in these two games: weak reliability from Player Bs (80.7% of decisions r) coupled with weak reliance from Player As (49% of decisions R) in BG1, and nearly universal reliability from Player Bs (94.3% of r) coupled with strong (yet imperfect) reliance from Player As (74% of R) in EG2.[6] Importantly, Jacquemet

[5] We avoid the end-game effect by providing no information about the exact number of repetitions.

[6] In the present study, we focus on the determinants on Player As' behavior, considering Player Bs' solely as a source of strategic uncertainty. Jacquemet and Zylbersztejn (2013) offer a systematic analysis of the patterns of Player Bs' decisions in this game.

Table 2: The experimental games

B

A	l	r
L	(9.75 ; 3.00)	(9.75 ; 3.00)
R	(3.00 ; 4.75)	(10.00 ; 5.00)

Baseline 1

B

A	l	r
L	(8.50 ; 3.00)	(8.50 ; 3.00)
R	(6.50 ; 4.75)	(10.00 ; 5.00)

Baseline 2

B

A	l	r
L	(9.75 ; 8.50)	(9.75 ; 8.50)
R	(3.00 ; 8.50)	(10.00 ; 10.00)

Egalitarian 1

B

A	l	r
L	(8.50; 8.50)	(8.50 ; 8.50)
R	(6.50 ; 8.50)	(10.00 ; 10.00)

Egalitarian 2

and Zylbersztejn (2014) provide systematic evidence that these outcomes cannot be explained by inequality aversion.

The payoff structures of BG1 and EG2 differ in terms of both players' monetary incentives to seek efficiency. In BG1, Player As may improve their situation only slightly when moving from L to (R, r) (from 9.75 to 10, a difference of .25), while a failed attempt to rely on Player Bs, resulting in (R, l), is very costly (yielding only 3 to Player As). In EG2, these two cases become more balanced – the gain for moving from L to (R, r) increases (from 8.5 to 10, a difference of 1.5), and the cost of relying on the other player in vain becomes less severe, with (R, l) now giving 6.5 to Player As. Analogous variations occur for Player Bs. The efficiency premium (conditional on Player As' reliance) is quite low in BG1 (from 4.75 to 5, a difference of 0.25) and drastically rises in EG2 (from 8.5 to 10, a difference of 1.5). Altogether, EG2 provides much more salient monetary incentives to act efficiently to both players.[7] In order to account for subjects' responses to the changes in their own as well as their partners' monetary incentives, we introduce two intermediate payoff matrices: Baseline Game 2 (BG2) and Egalitarian Game 1 (EG1) shown respectively in the top-right and bottom-left corners of Table 2. Each of the two games differs in only one dimension – that is, either Player As' payoffs or Player Bs' payoffs – as compared to BG1 and EG2: BG2

[7]Some studies have also documented the effect of the saliency of monetary incentives in coordination games. See, for example, Battalio, Samuelson, and Van Huyck (2001) for symmetric 2×2 games, and Goeree and Holt (2005); Devetag and Ortmann (2010) for n-player minimum and median effort games.

(EG1) has Player A's payoffs taken from EG2 (BG1) and Player B's payoffs taken from BG1 (EG2). Thus, the four games enable us to test the following hypotheses for the Human condition:

Hypothesis 1 *In the human treatment, the variations in the monetary incentives affect players' behavior as follows:*

(a) Player As react to the variations in their own monetary incentives to seek efficiency: the proportion of decisions R is higher in BG2 than in BG1, and in EG2 than in EG1;

(b) Player Bs react to the variations in their own monetary incentives to act efficiently: the proportion of decisions r is higher in EG1 than in BG1, and in EG2 than in BG2;

(c) Player As react to the variations in Player Bs' monetary incentives to act efficiently: the proportion of decisions R is higher in EG1 than in BG1, and in EG2 than in BG2.

The strategic uncertainty that Player As face cannot be directly observed or measured by the experimenter. Therefore, assessing its behavioral effect requires a benchmark in which the actual degree of strategic uncertainty can be controlled for. To that end, each of the four games is implemented under two different conditions: Human and Robot. In the Human treatment, two human subjects interact in ways described above. In the Robot treatment, a human subject acting as Player A interacts with a computerized Player B who is pre-programmed to always choose r. Subjects in the Robot treatment are clearly informed they are interacting with a pre-programmed computer: "**the computer chooses r at each round, without exception**" (bold in the original instruction sheet). This is the only difference in the rules and procedures between Human and Robot treatments.As a result, subjects in the Robot treatment do not face any strategic uncertainty. This leads us to the following hypothesis:

Hypothesis 2 *The Robot treatment neutralizes strategic uncertainty Player As face. As a result, the proportion of actions R in a given game is higher in the Robot treatments than in the Human treatments.*

While not choosing R in the Human treatment *may* not necessarily arise from strategic uncertainty, the same behavior in the Robot treatment *must* be due to reasons other than strategic uncertainty. Thus, comparing Player As' decisions between Human and Robot treatments enables us to capture the behavioral effect of strategic uncertainty.

Once the behavioral results are established, we then investigate the relationship between subjects' cognitive abilities and their sensitivity to the changes in the degree of strategic uncertainty. We measure each participants cognitive ability by implementing a standard psychological test developed by John C. Raven (see Raven, 2008, for an overview). Subjects are asked to solve Raven's test at the end of all experimental sessions involving the BG2 and EG1 games, under both Robot and Human treatments. Based on existing experimental results (discussed in the opening section), we formulate the following hypothesis:

Hypothesis 3 *Subjects with high cognitive abilities are more sensitive to changes in the degree of strategic uncertainty than those with low cognitive abilities.*

2.2 Experimental procedures

Upon arrival, participants are randomly assigned to their computers and asked to fill in a short personal questionnaire containing basic questions about their age, gender, education, etc. The pre-distributed written set of instructions is then read aloud. Player As are informed that they will play an unrevealed number of rounds of the same game, each round with a different partner, and that their own role will not change during the experiment. Before starting, subjects are asked to answer a quiz assessing their understanding of the game they are about to play. Once the quiz and any questions from participants are answered, the experiment begins.

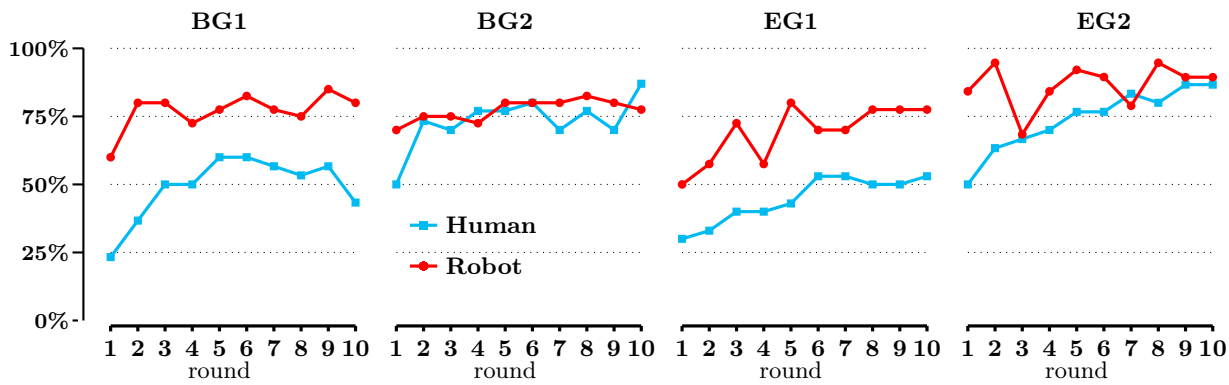
The experiment generates observations under eight experimental conditions, varying according to the payoff structure (BG1, BG2, EG1 or EG2) and Player A's partner (human subject in the Human treatment or computer in the Robot treatment). All conditions are implemented separately, using a between-subject design: each subject plays only one of the four games, and interacts either with other subjects or with a computer.

In addition, all sessions involving BG2 and EG1 include computerized post-experiment tasks. An additional 5 euros is paid to each subject for completing this part. Immediately after the end of the experimental game, participants are provided with a brief round-by-round summary of their decisions and outcomes, and are asked to provide any relevant comments and indicate the things that might have affected their decisions during the experiment in a blank space on their computer screens. They are also asked to solve a part of the advanced version of Raven's test – composed of 16 items to be completed within 10 minutes.

For each payoff matrix, we ran three Human treatment sessions (involving 20 subjects per session: 10 Player As interacting with 10 Player Bs), and two Robot treatment sessions (involving 20 Player As per session interacting with automated Player Bs). The data for the BG1 and EG2 Human treatments come from Jacquemet and Zylbersztejn (2014), while all the other sessions were carried out in October 2012 and February and March 2014. All sessions took place at the *Laboratoire d'Economie Experimentale de Paris* (LEEP) at University Paris 1 Panthéon-Sorbonne. Of the 398 participants (190 males), 323 were students with various fields of specialization.[8] The majority of subjects (57%) had already taken part in economic experiments run at LEEP. Participants' average age was 24.05. Subjects were recruited via an on-line registration system based on Orsee (Greiner, 2004) and the experiment was computerized through software developed under Regate (Zeiliger, 2000) and z-Tree (Fischbacher, 2007). Sessions lasted about 45-60 minutes, with

[8] In one EG2 Robot treatment session, we had 18 subjects instead of 20, so for the Robot treatment there were 40 subjects for BG1, BG2 and EG1, and 38 subjects for EG2. For the Human treatment sessions, we had 60 subjects (half of whom are Player As) for each of the four games.

Figure 1: Proportion of decisions R across rounds and treatments



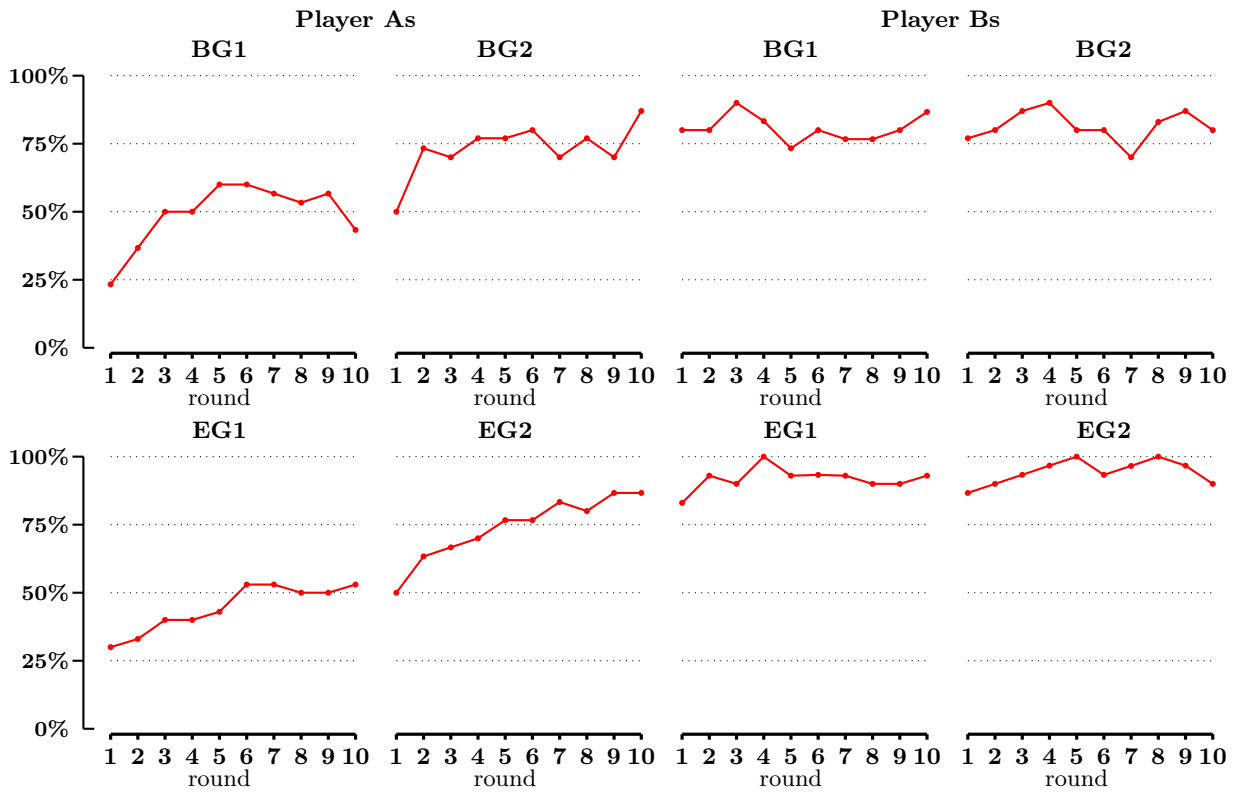
an average payoff of roughly 12.50 euros in Human treatments and 15 euros in Robot treatments (including a 5 euros show-up fee, but not the post-experiment task fee). No subject participated in more than one experimental session.

3 Results

Figure 1 provides descriptive statistics regarding the behavior of Player As in all our experimental treatments. The share of Player As who chose R in each round is displayed separately for each payoff matrix, and the two curves provide a comparison between the Human treatment and the Robot treatment. Before moving to a detailed analysis of the treatment effects, three main observations can be made. First, our Human treatment replicates the results seen in the existing literature: a high proportion of Player As decide to play L, even after several rounds of the game. Second, between game comparisons of behavior in the Human treatments show this pattern is barely influenced by the strategic context: while Player As react to changes in incentives they face (BG2 vs BG1, and EG2 vs EG1), they appear rather insensitive to changes in incentives faced by Player Bs (EG1 vs BG1, and EG2 vs BG2). Finally, while the share of decisions R in the Robot treatment always weakly dominates the one in the Human treatment, the absence of strategic uncertainty in this context does not remove all decisions L.

In the remainder of this section, the main question we seek to answer is whether and to what extent strategic uncertainty drives the observed decisions to play L. In the Human treatment, two factors explain the decisions of Player As: the behavior of Player Bs in the experiment and how Player As adjust to this behavior. The variations in payoffs between games allows these two factors to be measured separately. We then move to an analysis of the Robot treatments, in which strategic uncertainty is removed by design.

Figure 2: Share of the decisions R (r) for Player As (B) in the Human treatments, across rounds

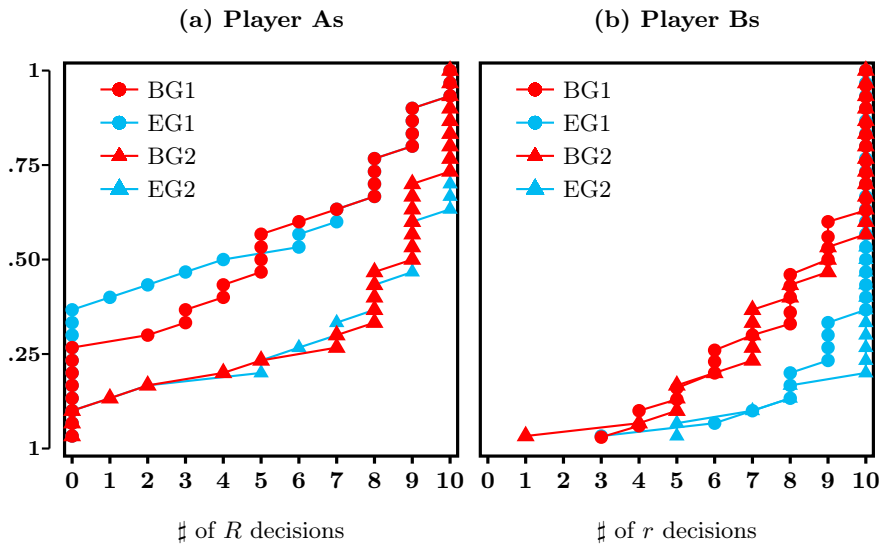


3.1 Results from the Human treatment

As summarized in Hypothesis 1, the variations in payoffs between games should induce variations in the decisions of both players, hence resulting in a variation in the actual strategic uncertainty faced by Player As. Figure 2 provides an overview of individual behavior in the Human treatment. Aggregate results suggest that both players react to the variations in their own payoff scheme. Holding Player Bs' payoffs constant, Player As are more likely to seek efficiency as their monetary incentives to do so become more salient: the frequency of R increases from 49% in BG1 to 73% in BG2 ($p = .001$), and from 45% in EG1 to 74% in EG2 ($p = .005$).[9] Analogously, Player Bs become

[9]We test the difference in proportion of a given outcome between two experimental conditions by carrying out a two-sided bootstrap proportion test that accounts for within-subject correlation – *i.e.* the fact that the same individual takes 10 decisions. The procedure consists of bootstrapping subjects and their corresponding decisions over all ten rounds instead of bootstrapping decisions as independent observations (see, *e.g.*, Jacquemet, Joule, Luchini, and Shogren, 2013, for a detailed description of the procedure). In Round 1, data are independent and thus allow us to analyze behavior with a standard bootstrap proportion test. Frequencies in Round 1 are 23.3% in BG1 and 50.0% in BG2 ($p = .027$), and 30.0% in EG1 and 50.0% in EG2 ($p = .091$).

Figure 3: EDF of the total number of decisions R and r decisions by game



more efficient the higher the cost of acting otherwise: the frequency of r increases from 81% in BG1 to 92% in EG1 ($p = .026$) and from 81% in BG2 to 94% in EG2 ($p = .012$). However, despite Player Bs' responsiveness to their personal monetary incentives, Player As remain insensitive to this factor: the differences between BG1 and EG1 and between BG2 and EG2 are small and not statistically significant ($p = .694$ and $p = .898$, respectively).[10]

These three observations are supported by the individual-level data summarized in Figure 3. The left-hand side provides the Empirical Distribution Function (EDF) of Player As' frequency of decisions R, *i.e.* the number of decisions R taken throughout the ten rounds of the game. Statistical tests indicate first order stochastic dominance of the distributions between BG2 and BG1 ($p = .007$) and between EG2 and EG1 ($p = .021$).[11] This confirms that Player As' behavior depends on the saliency of their own monetary incentives. Figure 3.(b) presents the EDFs of Player Bs' individual frequencies of decisions r. These individual-level data stand in line with the aggregate outcomes, showing that Player Bs do react to their personal costs of inefficient behavior: the EDF from EG1 first order stochastically dominates the EDF from BG1 ($p = .039$), and the EDF from EG2 first order stochastically dominates the EDF from BG2 ($p = .002$). This amounts to a significant change in the actual strategic uncertainty Player A faces. Figure 3.(a) also shows

[10]The analogous frequencies in Round 1 for Player Bs are 80.0% in BG1 and 83.3% in EG1 ($p = .731$), and 76.7% in BG2 and 86.7% in EG2 ($p = .232$); for Player As, they are: 23.3% in BG1 and 30.0% in EG1 ($p = .583$), and 50.0% in BG2 and 50.0% in EG2 ($p = .889$).

[11]Our first order stochastic dominance test is based on a bootstrap version of the univariate Kolmogorov-Smirnov (KS) test. The advantage of this test as compared to the standard KS test is it allows for ties (see, *e.g.*, Abadie, 2002; Sekhon, 2011).

that there are no significant differences between EG1 and BG1 ($p = .783$) or between EG2 and BG2 ($p = .914$); Player As do not seem to react to the changes in the saliency of Player Bs' monetary incentives.

Result 1 *In the presence of strategic uncertainty, the saliency of monetary incentives affects players' behavior as follows:*

(a) The proportion of decisions R is higher in BG2 than in BG1, and in EG2 than in EG1: Player As become more likely to act efficiently the more salient their monetary incentives;

(b) The proportion of decisions r is higher in EG1 than in BG1, and in EG2 than in BG2: Player Bs become more likely to act efficiently the more salient their monetary incentives;

(c) The proportion of decisions R is not significantly different in EG1 than in BG1, nor in EG2 than in BG2: Player As ignore the saliency of Player Bs' monetary incentives.

3.2 The effect of strategic uncertainty: comparing Human and Robot treatments

In the Robot treatment, Player Bs' decisions are generated by computers always playing r, hence eliminating the strategic uncertainty faced by Player As. The change in behavior from Player As hence identifies the effect of strategic uncertainty in the Human treatment. In Figure 1 we presented aggregate statistics on the behavior of Player As. Figure 4 provides individual data for all four games and both treatments. Each curve provides the EDF of the number of decisions R taken by each individual in the Human treatment and in the Robot treatment. In the aggregate, in three games out of four we find a statistically significant increase in the proportion of decisions R when Player As interact with computerized Player Bs rather than other human subjects: from 49.0% to 77.0% in BG1, from 44.7% to 69.0% in EG1 and from 77.0% to 86.6% in EG2. In BG2, however, we only observe a weak increase – from 72.5% to 77.3%.[12].

These results are true all along the distribution of the number of decisions R. The EDF in the Robot treatment first order stochastically dominates the EDF in the Human treatment for BG1 ($p < .001$), BG2 ($p = .004$) and EG1 ($p = .011$); this pattern is not statistically significant for EG2 ($p = .134$). The most noticeable difference between the two distributions is the proportion of Player As who invariably choose R when facing a computer rather than another subject: 50.0% against 10.0% in BG1, 65.5% against 30.0% in BG2, 37.5% against 10.0% in EG1 and 57.9% against 40.0% in EG2. This increase is significant at a 1% threshold for games BG1, BG2 and EG1 ($p = .001$, $p = .004$ and $p = .010$ respectively) and barely insignificant in EG2 ($p = .111$).

[12]These shifts are significant according to bootstrap proportion tests at the 1% threshold in BG1 ($p = .001$) and in EG1 ($p = .005$), at the 5% threshold in EG2 ($p = .033$) and are not statistically significant in BG2 ($p = .305$)

Figure 4: EDF of the individual number of decisions R by treatment

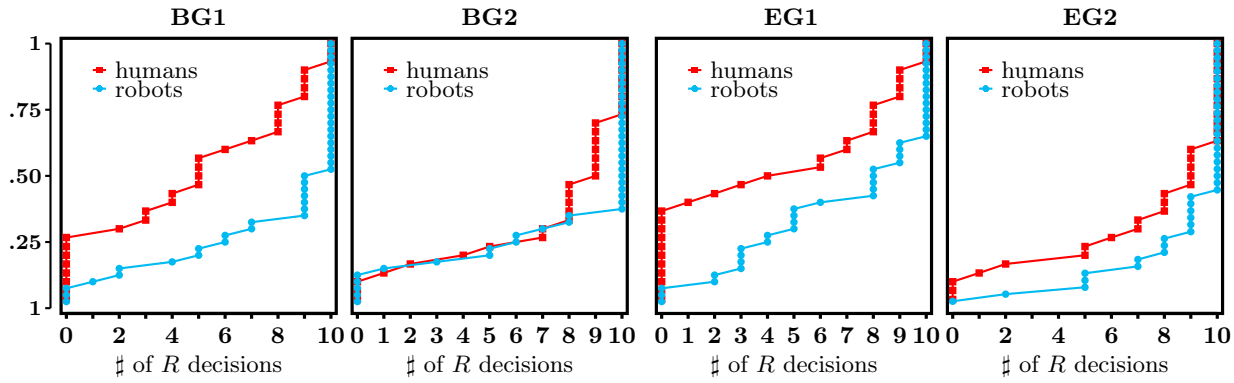


Table 3: Spearman's rank correlation of decisions R and Raven's test score

BG2	Human	0.055 ($p = 0.772$)
	Robot	0.571 ($p < 0.001$)
EG1	Human	-0.005 ($p = 0.978$)
	Robot	0.273 ($p = 0.088$)

Importantly, both aggregate and individual results unambiguously show that eliminating strategic uncertainty does not make Player As' behavior invariably efficient. Altogether, these observations lead us to the following result:

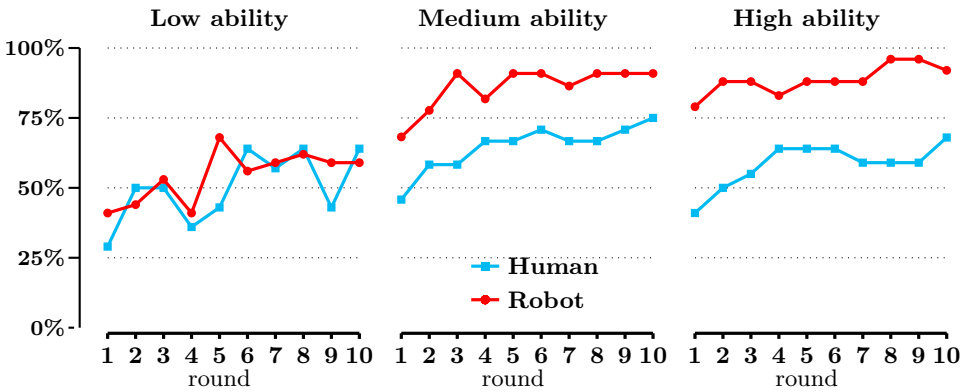
Result 2 *Strategic uncertainty only partially explains Player As' failure to seek efficiency in the coordination game. Player As do not consistently select the Pareto-Nash strategy R even when interacting with a computer who is known to always act efficiently by choosing r.*

3.3 Cognitive ability and strategic uncertainty

The crucial question raised by the Human-Robot comparison is why do Player As continue acting inefficiently in the absence of strategic uncertainty? We investigate this question based on the cognitive abilities of Player As. Cognitive ability is measured by Raven's test, performed after the game only in BG2 and EG1, in both Human and Robot treatments. It corresponds to the number of correct answers to the set of 16 questions.

The main outcome of interest is the change in the correlation of cognitive ability and actual decisions by treatment. The internal validity of the comparison depends on the randomization between treatments with regards to cognitive abilities. A multiple-treatment comparison using

Figure 5: Proportion of decisions R across rounds and treatments by ability group



Kruskal-Wallis test with Bonferroni correction does not reject the null hypothesis that Raven's test scores in the four experimental conditions come from the same population (with $p = .275$).[13] This result confirms that subjects are properly randomized regarding their cognitive abilities.

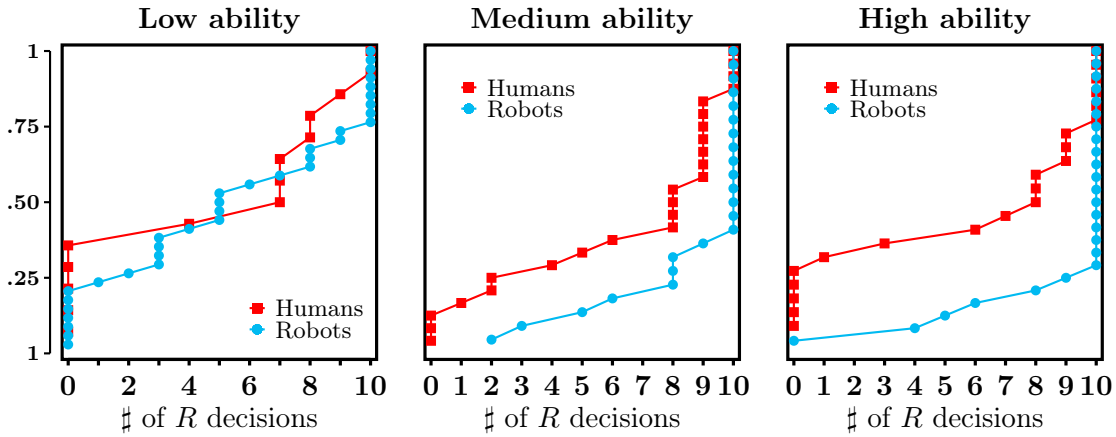
Table 3 reports the individual correlations between the score in Raven's test and decisions in the game, measured by the number of choices R (between 0 and 10) made throughout the 10 rounds of the game. We find a positive and significant correlation between these two variables in the Robot treatments which suggests that subjects with a higher cognitive ability are more likely to play R. However, this correlation disappears in Human treatments. To gain further insights on the transition between the Robot and Human treatments, we categorize individuals in our experimental sample according to their score on the Raven's test. Player As are divided into three groups, which correspond to the three tertiles of the overall score distribution, *(i.e.* considering all Player As from all four experimental conditions, hence 140 subjects.) Subjects in the low ability group have a Raven's test score below 7, those with a Raven's test score of 7 to 10 are in the medium ability group and those with a score above 10 are in the high ability group.

Figure 5 presents the proportion of decisions R across rounds by cognitive ability group.[14] The aggregate dynamics are very similar across cognitive ability groups, with an increase in the frequency of decisions R in the first rounds and a stabilization afterwards. This result holds regardless of whether decisions are taken with computers or humans acting as Player Bs. However, interacting with computers instead of humans induces an initial upward shift in the ratio of R that persists over time in the medium and the high ability groups, whereas no such shift occurs in the low ability group. The mean increase in proportion of decisions R in Round 1 induced by

[13]The same test applied at the session-level (10 sessions) instead of the experimental-condition-level yields a p-value of .694.

[14]The data from BG2 and EG1 are pooled to focus on the overall effect of removing strategic uncertainty and guarantee sufficient sample sizes in each category.

Figure 6: EDF of the total number of decision R by treatment and cognitive ability group



playing with a computer rather than a human is 22.4% ($p = .057$) for the medium ability group and 38.3% for the high ability group ($p = .003$). The mean difference between these proportions is still present in Round 10, but to a lesser extent: 15.9% ($p = .074$) for the medium ability group and 11.5% ($p = .022$) for the high ability group. In the low ability group, we find a small and insignificant upward shift in the proportion of decisions R in Round 1 (12.6%, $p = .203$), while the proportions of decisions R are almost equally likely in Round 10 (the difference being -5.5%, $p = .618$) in Robot and Human treatments.

Figure 6 provides the EDFs of the number of decisions R made by each subject across ten rounds for each cognitive ability group in Human and Robot treatments. Three results emerge. First, the EDFs in the Humans treatment are not statistically different across cognitive ability groups.[15] Second, this is not the case in the Robot treatment: the EDF for the low cognitive ability group is first order stochastically dominated by the EDF for the medium and high cognitive ability groups ($p = .003$ and $p = .001$, respectively). Finally, we observe no such relationship for the medium and high ability groups: the EDFs are not significantly different ($p = .318$). The main reason for these results is that the subjects in the low ability group do not respond to the absence of strategic uncertainty in Robot treatments by increasing the frequency of decisions R ($p = .415$), whereas subjects in the medium and high ability groups do so (the tests are based on comparisons between treatments for each cognitive ability group, the p-values for both are $p = .001$.)

Such behavior is in line with Hypothesis 3, *i.e.* Player As with higher cognitive abilities are more sensitive to the change in the degree of strategic uncertainty between Human and Robot

[15]To ensure a sufficient sample size in each ability group, we pooled the outcomes from both games in each treatment. The tests are performed using two-sided bootstrap K-S tests. The p-values of the two-by-two comparisons are: $p = .288$ for the low ability group vs. the medium ability group, $p = .599$ for the low vs. the high and $p = .695$ for the medium vs. the high.

treatments. To provide a formal test of the hypothesis, we first denote $\mathcal{R}$ the mean number of R decisions across 10 rounds by treatment (Robot vs. Human) and type (low cognitive ability vs. high cognitive ability). The test of Hypothesis 3 is then defined as follows:

$$\begin{cases} H_0 : \mathcal{R}(\text{Robot, high}) - \mathcal{R}(\text{Human, high}) = \mathcal{R}(\text{Robot, low}) - \mathcal{R}(\text{Human, low}) \\ H_1 : \mathcal{R}(\text{Robot, high}) - \mathcal{R}(\text{Human, high}) > \mathcal{R}(\text{Robot, low}) - \mathcal{R}(\text{Human, low}) \end{cases} \tag{1}$$

Empirical results indicate that our hypothesis is likely to be verified with $\mathcal{R}(\text{Robot, high}) = 8.8$, $\mathcal{R}(\text{Human, high}) = 5.8$, $\mathcal{R}(\text{Robot, low}) = 5.4$, and $\mathcal{R}(\text{Human, low}) = 5.0$, which is consistent with the hypothesis. We statistically tested the result by putting all the terms of Equation (1) to the left-hand side, so that the test reduces to a comparison test of multiple means with coefficients $(1; -1; -1; 1)$.[16] The test rejects the null hypothesis when the low cognitive ability group is compared to the high cognitive ability group ($p = .051$), as well as when it is compared to the medium and high cognitive ability groups pooled together ($p = .056$). These results allow us to state our last result:

Result 3 *Player As with high cognitive abilities are more sensitive to the changes in strategic uncertainty than those with low cognitive ability.*

4 Conclusion

Coordination failure is a widely documented phenomenon, with possibly dramatic economic consequences. While most experimental investigations try to identify the possible sources of strategic uncertainty underlying such failures, very few explore the effect of strategic uncertainty *per se*. In addition, little is known about whether and how behavior under strategic uncertainty is related to individual cognitive abilities.

We used four variations of a classic 2×2 dominance solvable coordination game to explore the link between coordination failure, strategic uncertainty and cognitive abilities. To isolate the behavioral effect of strategic uncertainty, we compare the decisions made when facing other human subjects (Human treatment) to those made against computer programs whose perfectly efficient behavior is common knowledge (Robot treatment). We find that the occurrence of coordination failure cannot be entirely explained by strategic uncertainty. First, a non-negligible share of our subjects failed to act efficiently even in the Robot treatment. Second, behavioral response to strategic uncertainty is related to individual cognitive abilities. Subjects with higher cognitive abilities (as measured by their scores on Raven's test) systematically adapt their decisions to the

[16] Given the group sizes, the bootstrap test is based on re-sampling subjects and the number of times they choose decision R. To account for asymmetry in the empirical distribution, we computed an equal-tail bootstrap p-value. See Davidson and MacKinnon (2006) for further details on this procedure.

varying degree of strategic uncertainty (i.e. between the Human and Robot treatment), while those with lower cognitive abilities fail to do so.

The relationship between cognitive abilities and the sensitivity to strategic uncertainty is a challenge for institutions aimed at fostering efficient coordination by reducing strategic uncertainty. These mechanisms – such as the widely studied affect of communication between participants – usually involve focusing agents' actions on the desirable equilibrium. However, the presence of players with a low strategic focus – such as the Raven's test low-scorers in our experiment – may undermine the effect and lower the economic value of such institutions. As noted by Ellingsen and Östling (2010) in the context of a theoretical analysis of communication and coordination in a level-k setup, "there is a need for evidence that systematically distinguishes the effects of preferences, belief, and rationality" (p.1714). We believe that our experimental design and collected data on the cognitive underpinnings of behavior under strategic uncertainty and the resulting coordination failure is an important step in this direction.

References

Abadie, A. (2002): "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Model," *Journal of the American Statistical Association*, 97(457), 284–292.

Akiyama, E., N. Hanaki, and R. Ishikawa (2013): "It is not just confusion! Strategic uncertainty in an experimental asset market," Aix-Marseille School of Economics, DP 2013-40.

Battalio, R., L. Samuelson, and J. Van Huyck (2001): "Optimization Incentives and Coordination Failure in Laboratory Stag Hunt Games," *Econometrica*, 69(3), 749–64.

Beard, T. R., and J. Beil, Richard O. (1994): "Do People Rely on the Self-Interested Maximization of Others? An Experimental Test," *Management Science*, 40(2), 252–262.

Beard, T. R., R. O. J. Beil, and Y. Mataga (2001): "Reliant behavior in the United States and Japan," *Economic Inquiry*, 39(2), 270–279.

Brañas-Garza, P., T. García-Muñoz, and R. Hernán (2012): "Cognitive effort in the Beauty Contest Game," *Journal of Economic Behavior and Organization*, 83(2), 254–260.

Brandts, J., D. J. Cooper, and R. A. Weber (2014): "Legitimacy, Communication and Leadership in the Turnaround Game," *Management Science*, forthcoming.

Burks, S. V., J. P. Carpenter, L. Goette, and A. Rustichini (2009): "Cognitive skills affect economic preferences, strategic behavior, and job attachment," *Proceedings of the National Academy of Science, U.S.A.*, 106(19), 7745–7750.

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

Carpenter, J., M. Graham, and J. Wolf (2013): "Cognitive ability and strategic sophistication," *Games and Economic Behavior*, 80, 115–130.

Cooper, D. J., and J. B. Van Huyck (2003): "Evidence on the equivalence of the strategic and extensive form representation of games," *Journal of Economic Theory*, 110(2), 290–308.

Cooper, R., and A. John (1988): "Coordinating Coordination Failures in Keynesian Models," *The Quarterly Journal of Economics*, 103(3), pp. 441–463.

Costa-Gomes, M. A., and V. P. Crawford (2006): "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," *American Economic Review*, 96(5), 1737–1768.

Davidson, R., and J. G. MacKinnon (2006): "Bootstrap methods in econometrics," in *Palgrave Handbooks of Econometrics*, ed. by T. C. Mills, and K. D. Patterson, vol. 1, chap. 23, pp. 812–838. Palgrave Macmillan, Basingstone.

Devetag, G., and A. Ortmann (2007): "When and why? A critical survey on coordination failure in the laboratory," *Experimental Economics*, 10(3), 331–344.

——— (2010): "Classic coordination failures revisited: the effects of deviation costs and loss avoidance," *Economics Bulletin*, 30(2), 1633–1641.

Ellingsen, T., and R. Östling (2010): "When does communication improve coordination?," *American Economic Review*, 100, 1695–1724.

Embrey, M. S., G. R. Fréchette, and S. F. Lehrer (2014): "Bargaining and Reputation: An Experiments on Bargaining in the Presence of Behavioral Types," *Review of Economic Studies*, Forthcoming, doi: 10.1093/restud/rdu029.

Fehr, E., and J.-R. Tyran (2001): "Does Money Illusion Matter?," *American Economic Review*, 91(5), 1239–1262.

Ferraro, P. J., and C. A. Vossler (2010): "The Source and Significance of Confusion in Public Goods Experiments," *B.E. Journal of Economic Analysis & Policy*, 10(1), 53.

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

Frederick, S. (2005): "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, 19, 25–42.

Gill, D., and V. Prowse (2013): "Cognitive abilities and learning to play equilibrium: A level-k analysis," Working Paper 641, Department of Economics, University of Oxford.

Goeree, J. K., and C. A. Holt (2001): "Ten Little Treasures of Game Theory and Ten Intuitive Contradictions," *American Economic Review*, 91(5), 1402–1422.

——— (2005): "An experimental study of costly coordination," *Games and Economic Behavior*, 51(2), 349–364.

Greiner, B. (2004): "An Online Recruitment System for Economic Experiments.," *University of Cologne, Working Paper Series in Economics*, 10, 79–93.

Heinemann, F. (2012): "Understanding Financial Crises : The Contribution of Experimental Economics," *Annales d'Economie et de Statistique*, 107-108, 7–29.

Houser, D., and R. Kurzban (2002): "Revisiting Kindness and Confusion in Public Goods Experiments," *American Economic Review*, 92, 1062–1069.

Ivanov, A., D. Levin, and M. Niederle (2010): "Can relaxation of beliefs rationalize the winner's curse? An experimental study," *Econometrica*, 78(4), 1435–1452.

Jacquemet, N., R.-V. Joule, S. Luchini, and J. Shogren (2013): "Preference Elicitation under Oath," *Journal of Environmental Economics and Management*, 65(1), 110–132.

Jacquemet, N., and A. Zylbersztejn (2013): "Learning, words and actions: experimental evidence on coordination-improving information," *B.E. Journal of Theoretical Economics*, 13(1), 215–247.

Jacquemet, N., and A. Zylbersztejn (2014): "What drives failure to maximize payoffs in the lab? A test of the inequality aversion hypothesis," *Review of Economic Design*, 18(4), 243–264.

Johnson, E. J., C. Camerer, S. Sen, and T. Rymon (2002): "Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining," *Journal of Economic Theory*, 104, 16–47.

Morris, S., and H. S. Shin (2003): "Global Games: Theory and Applications," in *Advances in Economics and Econometrics*, ed. by M. Dewatripont, L. Hansen, and S. Turnovsky, chap. 3. Cambridge University Press.

Mullainathan, S., and E. Shafir (2013): *Scarcity: Why Having Too Little Means So Much*. Times Books, Henry Holt and Company, LLC, New York, NY.

Polonio, L., S. Di Guida, and G. Coricelli (2014): "Strategic sophistication and attention in games: an eye-tracking study," ECARES working paper 2014-22, ECARES.

Raven, J. (2008): "General Introduction and Overview: The Raven Progressive Matrices Tests: Their Theoretical Basis and Measurement Model," in *Uses and abuses of intelligence*, ed. by John, and J. Raven, chap. 1, pp. 17–68. Competency Motivation Project, Edinburgh, Scotland.

Rosenthal, R. W. (1981): "Games of perfect information, predatory pricing and the chain-store paradox," *Journal of Economic Theory*, 25(1), 92–100.

Sekhon, J. (2011): "Multivariate and Propensity Score Matching Software with Automated Balance Optimization," *Journal of Statistical Software*, 42(7), 1–52.

Selten, R. (1975): "Reexamination of the perfectness concept for equilibrium points in extensive games," *International Journal of Game Theory*, 4(1), 25–55.

Smith, V. L., G. L. Suchanek, and A. W. Williams (1988): "Bubbles, Crashes. and Endogenous Expectations in Experimental Spot Asset Markets," *Econometrica*, 56(5), 1119–1151.

Zeiliger, R. (2000): "A presentation of Regate, Internet based Software for Experimental Economics," *http://regate-ng.gate.cnrs.fr/ sferriol/*.