DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

GAME PERCEPTION AND HARMONY IN 3×3 GAMES: AN EXPERIMENT IN COGNITIVE GAME THEORY

Daniel John Zizzo

Number 118

September 2002

http://www.economics.ox.ac.uk

Manor Road Building, Oxford OX1 3UQ

Game Perception and Harmony in 3×3 Games: An Experiment in Cognitive Game Theory

Daniel John Zizzo*
Department of Economics and Christ Church College
University of Oxford

(Version September 3, 2002)

Abstract

The experiment presented in this paper employs 3×3 games to analyze how perception of a game affects behavior in the presence or absence of a minimal framing effect and of uncertainty about the values of some game payoffs. We vary the harmony of practice stage games, and explain how this changes later behavior. We employ techniques, such as payoff integration and similarity evaluations, that could be used in further research to open the black box of framing effects. Game harmony is a measure summarizing how harmonious the interests of the players are in the game. It is associated with cooperation.

JEL Classification Codes: C72, C91, D83.
Keywords: 3×3 games, game harmony, cooperation, cognitive game theory.

* Email address: daniel.zizzo@economics.ox.ac.uk. I would like to thank Abigail Barr for help, Jim Engle-Warnick and seminar participants in Oxford and Royal Holloway for advice, and Jahirul Islam and Jonathan Tan for experimental assistance. The usual disclaimer applies. Financial support from the George Webb Medley Fund, Oxford, is gratefully acknowledged. This paper is dedicated to the memory of Michael Bacharach.

1. Introduction

This paper is based on the intuition that how an economic agent perceives a game is important for her to decide how to play it. Whether, for example, an agent perceives a Prisoner's Dilemma as such or as a Chicken or perhaps even a coordination game may make a difference in whether she decides to cooperate or not. Framing effects can have pervasive, and yet poorly understood, effects on cooperation rates (e.g. Cookson, 2000): we argue that framing effects can be understood as being due to differences in the way that decision problems are perceived, and provide tools to make progress in understanding how they operate. Further, one can shed light on how certain perceived game features, such as 'game harmony', affects cooperation in games.

Game harmony is identified as a scalar measure summarizing how harmonious or disharmonious the interests of the players are in the game (Zizzo, 2002). A pure coordination game is an example of perfectly harmonious game, and a zero-sum game one of a perfectly disharmonious game: most games are somewhere in the middle.

Our experiment finds a strong correlation between game harmony and cooperation in games. In addition, we employ the same "minimal framing" manipulation employed with 2×2 games by Zizzo and Tan (2002): practice stage games can be of different average levels of *game harmony* across games (high, medium or low). Hence, in the practice stage subjects faced games entailing different levels of "cooperativeness". The simplest framing effect prediction is that subjects should play more cooperatively in later game play the more harmonious the games they have faced in the practice stage.

Agents often face incomplete and ambiguous decision problems in economic transactions. This incompleteness and ambiguity may derive from a variety of sources, from lack of information to contract incompleteness to cognitive limitations in processing information. If there is any residual of incompleteness in the specification of a game for any reason, agents will have to form their own mental models of the decision problem and, in so doing, they may use any clue they can get. The clues may be provided by *game antecedents*, i.e. by anything that may have happened before the game. For example, subjects may have been asked to do some task inducing a 'cooperative frame' or an 'entrepreneurial frame': much experimental evidence suggests that this or other framing effects have an impact on cooperation (e.g., Elliott et al., 1998; Morris et al., 1998; Cookson, 2000). Framing effects might be thought to operate more the more the ambiguity about the game. Yet, game ambiguity may affect game play also in another way: by increasing its complexity, behavior in it might be better describable by simpler, less cognitively demanding algorithms. Our experiment uses 3×3 games in order to enable or facilitate the identifiability of play according to different non-cooperative algorithms. For example, if we had 2×2 games one would not be able to distinguish D1 (defined as one round of strict iterated dominance) from D2 (defined as two rounds of strict iterated dominance).

We develop methods to try to understand better how subjects deal with uncertainty about the game payoff values and how this interacts with framing effects: these methods are based on the incentive-compatible elicitation of payoff guesses and of similarity evaluations between games. We believe that the development of tools to understand framing effects is an important task, since all too often framing effects are noted but not explained, notwithstanding their potentially large effect on behavior (for example, on cooperation in social dilemmas).

The rest of this paper is organized as follows. Section 2 describes the concept of game harmony and otherwise presents the experimental background. Sections 3, 4 and 5 present the experimental design, hypotheses and results, respectively. Section 6 concludes. Appendices A and B contain the experimental instructions and a list of regression variables (as referred to in Section 5.3), respectively.

2. Experimental Background

This section briefly clarifies the concept of game harmony as discussed in Zizzo (2002) and reviews the evidence on game harmony and game perception.

2.1 Game harmony

Following Zizzo (2002), let Γ be a finite n-person game in normal form, and let N be the set of players such that $|n| = N$. Denote W_i the actions available to player i, so $W = W_1 \times W_2 \times ... \times W_n$ is the set of possible outcomes or states of the world, each of which we label by w in W. Payoffs are defined by x_{iw}: $W \rightarrow \mathbb{R}$, a standard Von Neumann-Morgernstern utility function, and so x_{iw} is the payoff for player i ($i \in N$) in state of the world w. For n players, there are $C = ½\, n\, (n - 1)$ player pairs. Let us label the payoffs x_{iw} for each pair c as a_{cw}, b_{cw} for $w \in W$. We can then define the cardinal harmony G(Γ) of game Γ as the arithmetic mean of the Pearson correlations between the C pairs of Γ:

$$G(\Gamma) = G(x_i, W) = \frac{1}{C}\sum_{c=1}^{C} r_c(a_{cw}, b_{cw}) = \frac{1}{C}\sum_{c=1}^{C}\frac{Cov(a_{cw}, b_{cw})}{\sigma_a \sigma_b} \qquad (1)$$

In the case of 3 × 3 games, $C = 1$, so G(Γ) = r (a_w, b_w), which is obviously bounded between -1 and +1. Zizzo (2002) proves that G(Γ) is scale-independent and is also invariant to transformations of Γ into a game Ψ that is the aggregation of any number of replicas of Γ, so G(Γ) = G(Ψ). An equivalence result can be shown to hold between perception of a game as more or less harmonious and psychological payoff transformation. Let us consider a player i as having interdependent preferences in relation to another player j if her utility function can be written as:

$$V_{iw} = x_{iw} + \sum_{j \neq i}^{n} \beta_i v(x_{jw}) \qquad (2)$$

where $v(x_{jw})$ is a positive monotonic transformation of x_{jw} and $\beta_i \in [-1,1]$ is the weight on j's utility component. Denote Γ^u the untransformed payoff matrix and Γ^v the payoff matrix transformed according to V. One can then prove that, if one takes the material payoffs x_{iw} and x_{jw} as given, $G(\Gamma^v) \gtreqless G(\Gamma^u) \leftrightarrow \beta_i \gtreqless 0 \ \forall i$, and $\partial\beta_i / \partial G(\Gamma^v) > 0$. This "equivalence result" implies that, if for whatever reason (e.g., framing or a personality trait) the agent perceives the game as more harmonious, she will behave as if she had more positive, or less envious, preferences. This will entail greater cooperation in all games where such simple payoff transformation will help attaining it (such as in the Prisoner's Dilemma, PD in what follows).

A closely related measure of game harmony can be obtained by considering payoff orderings rather than cardinal values. Let X_i be the set of all payoff values for player i in W, and let $x_{iw}^{\rho} = rank(x_{iw} | X_i)$, which can be mapped into rank payoff pairs $a_{cw}^{\rho}, b_{cw}^{\rho}$. Then:

$$G_\rho(\Gamma) = G_\rho(x_i, W) = \frac{1}{C}\sum_{c=1}^{C} r_c(a_{cw}^{\rho}, b_{cw}^{\rho}) \qquad (3)$$

In the case of 2 × 2 games, this reduces to $G_\rho(\Gamma) = r_c\,(a_{cw}^{\rho}, b_{cw}^{\rho})$. While undoubtedly not the only possibly measures of game harmony, G(Γ) and $G_\rho(\Gamma)$ have the virtues of general applicability, simplicity and lack of degrees of freedom.

Zizzo and Tan (2002) found that these measures could explain some two thirds of the variance in mean cooperation rate in their sample of mostly well-known 2×2 games, such as the PD, the Stag-Hunt, the Chicken, a coordination game and three variants of trust games. They also had the minimal framing manipulation discussed in the introduction and found that it was ineffective in changing the mean cooperation rate.

2.2 Game perception

The experiment presented in this paper differs from that of Zizzo and Tan (2002) in at least two important respects. First, the game samples are different: we employ 3×3 games and, with two exceptions (two 3×3 variants of the PD), we use randomly generated games throughout the experiment. Second, while in Zizzo and Tan (2002) payoff matrices are always displayed in their complete form, our experiment has treatments where some payoff cells are hidden and in relation to which subjects are asked to guess incentive-compatibly what the payoff values are, before playing each game. There are also treatments where some payoff cells are hidden but there is no guessing task or where, conversely, subjects are simply asked to guess random numbers.

There has been only limited work in experimental economics trying to address how agents perceive the game they are playing. In Mitropoulos' (2001) experiment, subjects were unaware about the structure of the game except that payoffs cells were either 0 or 1, and unaware of the actions by the other player: they were actually playing a very simple 2×2 game (the mutual fate control game), and Mitropoulos was interested in how subjects got to coordinate on the equilibrium. Costa-Gomes et al. (2001) worked experimentally with a variety of normal-form games: they collected useful information on the cognitive process guiding agents to action by allowing only one payoff cell to be open for the subjects to see at any one time, with a click of the mouse. They found most support for L1 and L2, rather than for the Nash algorithm. L1 corresponds to choosing the best action against a uniformly randomizing opponent; L2 to choosing the best action against a L1 player. Oechssler and Schipper (2000) were also interested in how agents learn to know about the structure of the game: subjects know their normal form game matrix payoffs, but not those of the other player. They could infer information about the structure of the game from the feedback on the actual payoff received after each round. They were also asked, and knew they would be asked, questions on the payoff structure (ordering between payoffs), so that there was an actual incentive for guessing the correct structure of the game. One interesting aspect of these papers is that cognitive processes are studied by *hiding away*, one way or another, information about the payoff structure at any given time. We follow the lead, but expand it in a new direction by asking subjects directly to make guesses on the payoff values.

3. Experimental Procedure

3.1 Design overview

The experimental sessions were run in the Department of Economics in the University of Oxford in May 2001 (Treatments 4-7), December 2001 (Treatment 8) and May 2002 (Treatments 1-3). Treatments 1 through 7 had 16 subjects each, for a total of 112 subjects;[1] treatment 8 was a control task at the start of a different experiment involving 61 subjects. Subjects were mostly undergraduate or postgraduate students, but not exclusively so. All treatments were computerized and partitions were used between subjects to ensure the anonymity of their decisions. An overview of the treatments is provided by Table 1.

(Insert Table 1 about here).

Treatments 1 through 7 had a common structure. Four subjects participated to each session. There were three stages.[2]

In Stage 1, subjects did practice by choosing actions in fully displayed 3×3 games, one game at a time in six rounds. Subjects were matched with the same coplayer throughout the stage, and received feedback on game outcomes after each round. The practice games set varied across treatments, as analyzed shortly.

In Stage 2, subjects were presented with ten 3×3 games in sequence, and never received any feedback on the outcome of their choices after each game in the sequence. They knew they were matched with a different coplayer from the one of the practice stage. What subjects saw on the computer screen and had to do varied across treatments, but, whatever the treatment, subject always had to choose an action in relation to the game.

[1] We later discovered that in Treatment 7 a subject was a (heavily game-theory-trained) Oxford M.Phil. Economics student, and another one had done the Zizzo and Tan (2002) experiment just a few weeks earlier. Their choices are removed from all the analysis below (and from the sample size as reported in Table 1).

[2] In order to check understanding of the instructions, subjects filled questionnaires at the start of each stage. Their answers were checked by experimenters, and, if any was incorrect or missing, the relevant points were explained individually.

Stage 3 involved no interaction and this was known in advance: subjects were simply to ask to assess incentive-compatibly the similarity of sixteen game matrices presented in sequence on the computer screen to a comparison game they had in print (the "comparison decision table", CDT for short). The focus of this paper is on Stage 2, but we shall refer briefly to some result from Stage 3.

3.2 Stage 1 (Treatments 1-7)

There were three different experimental conditions in Stage 1, according to whether subjects faced a sample of games with mean high, medium or low game harmony (High, Medium and Low condition, respectively), as measured by $G(\Gamma)$ and $G_\rho(\Gamma)$: Low corresponded to Treatments 1 and 4, Medium to Treatments 2, 5 and 7, and High to Treatments 3 and 6 (see Table 1).

Games were chosen drawing payoff values randomly from a uniform distribution between 0 and 100, under the restriction that they all had to have a unique pure Nash equilibrium; this was to guarantee that games would be of not too widely different strategic complexity.[3] Mean $G(\Gamma)$ was -0.890, 0.159 and 0.902 in the Low, Medium and High condition, respectively: Table 2 lists the games used.

(Insert Table 2 about here).

Subjects played for six rounds with the same player. Each round they faced a different game, and after the decisions were made they would receive feedback on the game outcome. Subjects did not play the practice stage for real money.

3.3 Stage 2 (Treatments 1-7)

Subjects played ten different games, this time not receiving any feedback on the game outcome. They were matched with a different coplayer than the one they had faced in Stage 1. They also knew that this was the last stage in which they would interact with other participants.

Treatments 1-3 (Visible condition). In Treatments 1-3, subjects simply had to choose an action each round. Game matrices were fully visible on the computer display. They were games with randomly generated payoffs, except those in rounds 9 and 10, which were variants on the PD.

(Insert Table 3 about here).

[3] We also checked that the game samples were not otherwise meaningfully different in other undesired respects (e.g., in the distribution of mean payoff values or in that of equilibria according to non-Nash algorithms).

In the payment stage at the end of the experiment, a round would be picked up randomly by the computer to determine the *action payment* on the basis of the action payoff points earned in that round. The resulting payoffs were converted in UK pounds at the rate of £ 0.05 per experimental point.

Treatments 4-6 (Guess condition). In Treatments 4-6, the right diagonal of the game matrix was "hidden" by question marks (see Figure 1).

(Insert Figure 1 about here).

Subjects were first asked incentive-compatibly to replace the question marks with their best guesses of the expected payoffs ("to replace the question marks with the point numbers that you guess correspond to each hidden cell") and then they had to choose actions. The guesses were just such, and had no effect on the actual payoff values hidden in each cell. Subjects knew that payoffs had to be between 0 and 100. At the end of the experiment, the *guess payment* was determined on the basis of the guess made in a randomly chosen round, different from the one determining the action payment. In relation to this round, subjects were paid £ 3 to get the guess exactly right, with a penalty of one £ 0.12 per each point of error (subjects were paid zero for getting the guess wrong by 25 points or more). Thus, a standard absolute difference incentive-compatibility mechanism was implemented (e.g., Croson, 2000). The instructions contained a table with details on the amount of payment for any given level of error, and so subjects were not required to make any significant computation.

Treatment 7 (Frozen condition). Treatments 1 through 6 cross two factors: practice stage game harmony and need for payoff integration. However, the Guess condition differs from the Visible condition not in one but in two ways: subjects face the payoff cells covered by the question marks (the "hidden cells") *and* they have to guess what payoff values are hidden by the question marks. We had a control condition that had the first change but not the second: in Treatment 7, the right diagonal of the game payoff matrix was hidden by question marks, but these were "frozen", i.e. they could not be replaced with numbers. Subjects simply had to choose game action on the basis of the incompletely displayed payoff matrix.

3.4 Stage 3 (Treatments 1-7)

Subjects were asked to assess incentive-compatibly the similarity of sixteen game matrices presented in sequence on the computer screen to another game, the "comparison decision table" (CDT in what follows, see Figure 2), they had in print.[4]

(Insert Figure 2 about here).

The games were different from those of Stage 2, and the payoffs were chosen drawing numbers between 0 and 100 from a uniform distribution, with eight games out of the sixteen having unique Nash equilibria. The order of presentation of the games was randomized across subjects. Another incentive-compatibility mechanism was in place to determine the *similarity payment*: one round out of sixteen was chosen randomly, and an absolute difference incentive-compatibility mechanism was in place, with £ 6 as maximum payment, and £ 0.24 deducted for every point off the mark. Payment was determined at the end of the experiment and subjects received no feedback on the outcome of their choices during Stage 3. For this reason, the determination of the "correct" similarity answer was an issue that had to be practically addressed in order to ensure financial motivation and determine payments, but not one with serious bearing on the experiment. This was important since any choice of "correct" similarity values was bound to be somewhat arbitrary, and a potential source of distortions in the experiment if learning feedback had been provided.[5] In total, payments could range between £ 4 and (depending on the treatment) £ 15 or 18; actual mean payments were in the order of 8-9 pounds for about 45-75 minutes of work (a short questionnaire was administered at the end).

As the Stage 3 games were different from those of Stage 2, the similarity choices will be only limitedly useful for the analysis of this paper. Nevertheless, we shall mention one result from Stage 3 that is directly relevant to how Stage 2 games were perceived.

3.5 Treatment 8

---

[4] This methodology is well-known in cognitive psychology (e.g., Hahn and Chater's, 1997, review). In economics, Buschena and Zilberman (1999) asked subjects to evaluate the similarity between standard binary lottery choices. Gilboa and Schmeidler's (2001) case-based decision theory is a possible formalization of reasoning by similarity.

[5] The "correct" similarity value was determined giving equal weight to the absence of a unique pure Nash equilibrium (as in both the CDT) and to the Euclidean distance between the payoffs (*EDP*). Let *Unique* be equal to 1 when there is a unique pure Nash equilibrium, and let *max*(*EDP*) and *min*(*EDP*) the maximum and minimum *EDP* between a game among the sixteen and a CDT. Then the "correct" answer was determined as $50 \times Unique + 50 \times [(EDP - min(EDP)] / [max(EDP) - min(EDP)]$.

Treatment 8 was an individual choice experiment, and this was known in advance. It worked as a control condition with respect to the guessing task of Treatments 4 through 6. Between 13 and 18 subjects had to guess a number between 0 and 100 in each of six cells. One of the six cells was then chosen at random. The number chosen in this cell was compared by the computer to a randomly determined "correct" answer and subjects were paid according to the same absolute difference payment scheme used in Treatments 4-6: they got a guess payment equal to 3 UK pounds minus 12 pence for every point by which the guess was incorrect (with a minimum of 0 if the guess was wrong by 25 points or more).

4. Experimental Hypotheses

We can now define our experimental hypotheses with more precision; all hypotheses are defined in relation to Stage 2 games when not specified otherwise.

H1 (Game harmony). Higher game harmony is associated to higher mean cooperation rates (as in Zizzo and Tan, 2002). In Treatments 1 through 3, where subjects are faced with complete payoff matrices and hence can compute the harmony of the game, we should expect a positive correlation between game harmony and mean cooperation rate.

H2 (Assimilation effect). The framing effect induced by the game harmony in the practice stage operates by *assimilation* to the harmony of the practice stage games. If so, mean cooperation is greater, and in the Guess condition δ (as defined below) is smaller, the higher the harmony of the practice stage games.

A framing effect could also operate by *contrast* of a game with a previous game: for example, an agent who has had experience with very harmonious games might notice how different, in terms of harmony, those games appear relative to the game she is now playing, and so, by contrast, cooperate less than she otherwise would.[6] Any contrast effect works in the opposite direction to the assimilation effect, and the two could cancel each other out. It may be difficult to discriminate between the absence

[6] Contrast effects are well known in psychophysics (see Kahneman and Varey, 1991).

of framing effect and existence of a contrast effect. One way of doing so is to look at the guesses in Stage 2: define a variable, δ, as the root mean square of the difference between one's own guess of one's own and the coplayer's gains - a_{ij}^{g} and b_{ij}^{g} (respectively) -, averaged out across the three right diagonal outcomes:

$$\delta = \left[\frac{(a_{13}^{g} - b_{13}^{g})^2 + (a_{22}^{g} - b_{22}^{g})^2 + (a_{31}^{g} - b_{31}^{g})^2}{3} \right]^{\frac{1}{2}} \tag{4}$$

δ measures how different a subject guesses her payoff relative to that of her coplayer. If there is an assimilation effect, we would expect δ to be smaller the greater the harmony of the practice stage games. This would not happen if a contrast effect were to dominate. An alternative possible way is by analyzing Stage 3 similarity choices: if a subject is being liable to a contrast effect, then we would expect a significantly lower similarity rating between games than if (only) an assimilation effect were operative.

H3 (Contrast effect). The framing effect induced by game harmony in the practice stage also operates by *contrast* to the harmony of the practice stage games. It would be evidence for this if we found that a) δ does not vary as predicted by H2 and b) the mean similarity rating in Stage 3 is lower while the Stage 2 cooperation rate is lower or anyway not higher.

H4 (Interaction effect). If we accept that the Guess condition presents more ambiguous decision problems to the subjects than those in the Visible condition, then subjects might rely more on past experience, i.e. we would expect a greater effect of practice stage game harmony on Stage 2 cooperation in the Guess condition than in the Visible condition.

Whether *H4* holds or not will depend on the counterbalancing of assimilation and contrast effects. The last two hypotheses are related to how the guesses are formed and to how the problem ambiguity affects decision-making.

H5 (Guesses as non-random). Guesses give an indication of the expected payoff values by the subjects in the hidden cells of Treatments 4-6. Thus: 1) subjects do not choose payoff guess values at random; 2) they take guesses into account in choosing what action to play.

H6 (Problem ambiguity). Under the assumption that H5 holds, subjects appear to rely on simpler algorithms when dealing with incomplete payoff matrices than otherwise. The cooperation rate may also change.

5. Experimental Results

We shall organize the presentation of our data around the experimental hypotheses listed in the previous section.

5.1 Test of H1: Cooperative Outcomes and Game Harmony

In order to test for the association between $G(\Gamma)$ [or $G_\rho(\Gamma)$] and cooperation, we need to define what we mean by a "cooperative action": while this is uncontroversial in the two variants of the PD, it may not be so in the other games. We define a cooperative action as the action associated to the utilitarian solution, i.e. the combination of actions yielding the highest sum of payoffs.[7] We focus on Treatments 1-3, not just because they are the only ones where subjects had full knowledge of the material payoff matrix: after all, in Treatments 4-6 we could rely on the subjective guesses by the subjects to "complete" the payoff matrix and obtain a subjective ("guess-integrated") payoff matrix. We use this method in the next subsection. However, it is inapplicable here since no two guess-integrated payoff matrices are exactly the same, and so we are not able to estimate a mean cooperation rate, for each guess-integrated payoff matrix, over a sample sufficiently large to be informative.

Let us define c_g as the mean cooperation rate in each game in Treatments 1-3, those where subjects had full knowledge of the material payoff matrix. H1 predicts a positive correlation between $G(\Gamma)$ [or $G_\rho(\Gamma)$] and c_g: this is indeed the case, since Pearson $r[G(\Gamma), c_g] = 0.627$ and Spearman $\rho[G(\Gamma), c_g] = 0.579$, while Pearson $r[G_\rho(\Gamma), c_g] = 0.601$ and Spearman $\rho[G_\rho(\Gamma), c_g] = 0.603$ (in both cases, $P < 0.05$, $n = 10$). This correlation is illustrated in Figure 3, where game 4 is the noteworthy outlier.

[7] An alternative that gives similar results is to employ the Nash Bargaining solution, i.e. to refer to the combination of actions yielding the highest *product* of payoffs.

(Insert Figure 3 about here).

It is striking to find that the cooperation rates in the PD variants were below 10% in our experiment, definitely below those found in psychological experiments (e.g., Pruitt and Kimmel, 1977). Since the game is a 3 × 3 variant of the PD, it is possible for the agent to cooperate *partially*, and this is not captured by c_g. The fact that c_g is a less adequate measure with 3 × 3 games than it is with 2 × 2 games may explain why our correlation coefficient, while large, is significantly lower than the one (about 0.9) found by Zizzo and Tan (2002).

A more detailed analysis would recognize that, due to the asymmetric nature of all the games (with the exception of the PD variants), we should really analyze the cooperation rate by player role, rather than just by game. This is bound to decrease the correlation coefficient, since this within-game cooperation rate variance cannot be explained by a measure that applies at the level of the game rather than at that of the single player.[8] Let c_r be the mean cooperation rate by player role rather than just by game. The correlation coefficients drop by about 0.1-0.15, which is comparable to the equivalent analysis in Zizzo and Tan (2002): $r[G(\Gamma), c_r] = 0.485$ and Spearman $\rho[G(\Gamma), c_r] = 0.456$; $[G(\Gamma), c_r] = 0.466$ and Spearman $\rho[G(\Gamma), c_r] = 0.500$ (in all cases, $P < 0.05$, $n = 20$).

Notwithstanding the qualifications, there is strong support for *H1*:

RESULT 1. *Simple game harmony measures are strongly associated with cooperation, as predicted by H1.*

5.2 Test of H2-H4: Framing and Game Perception

The mean cooperation rates by treatments c_t are listed in Table 1. In the Frozen condition, the cooperative outcome was defined over the visible payoff cells. In the Guess condition, it was defined over the guess-integrated payoff matrix, the game matrix that replaces the question marks with the guesses by the subject.[9]

c_t generally increases with the practice stage game harmony, as for *H2*, but with one exception: in the Guess condition, Low and Medium (i.e., Treatments 4 and 5) both have $c_t \approx 0.47$. A *F* test

[8] One could devise less parsimonious measures of game harmony that are subject-specific: see Zizzo (2002) for details.

[9] This method implicitly relies on *H5*, the evidence for which is discussed in section 5.3.

employing game harmony condition H and game matrix incompleteness condition I as factors yields significance for I [F (1, 1093) = 5.921, $P < 0.01$], marginal significance for H [F (2, 1093) = 2.791, $P < 0.07$], and insignificance for the interaction effect [F (2, 1093) = 0.594, *n.s.*]. By lumping Low and Medium together, we get significance not only on I [F (2, 1095) = 6.121, P < 0.01] but also on H [F (2, 1095) = 5.359, P < 0.05]; the interaction effect remains insignificant [F (1, 1095) = 0.780, *n.s.*].

How do we reconcile our finding of marginal significance of H with the insignificance with 2×2 games pointed out by Zizzo and Tan (2002)? The answer comes from doing separate F tests on H in the Visible condition and in the Guess condition: in the Visible condition, F (2, 467) = 0.827 (P = 0.438, *n.s.*); in the Guess condition, F (2, 477) = 2.570 (P = 0.078). Thus, even though there is not a significant interaction effect with the more global F tests, the marginal significance of H in them appears driven by the Guess treatments.

By lumping Low and Medium we are increasing the significance of H since we are neglecting that c_t is the same in Treatments 4 and 5. The existence of a counterbalancing contrast effect, as for *H3*, may explain this finding. To verify whether this is so, we now look at the Stage 2 δ values and at the Stage 3 similarity choices.

Table 1 lists the mean δ and similarity values by treatment s_t. A F test on δ using H as factor is significant [F (2, 477) = 21.970, $P < 0.001$]: Scheffe' and Tukey post-hoc tests indicate a significance of the difference between High and the other two conditions ($P \leq 0.001$). While this supports *H2*, the finding that δ is lower, or at least the same, in Low relative to High is a strong indication that a contrast effect is at work (as for *H3*).

The lowest similarity value is for Treatment 4, the Low, Guess condition. A F test of I and H on s_t yields insignificance for I [F (2, 1753) = 0.585, *n.s.*], significance for H [F (2, 1753) = 3.427, $P < 0.05$] and more so for the interaction effect [F (2, 1753) = 14.844, $P < 0.001$]. In the light of the significance of the interaction effect, there is therefore evidence that subjects tend to evaluate games as less similar to one another in Treatment 4, as a contrast effect would predict.

We can summarize our results as follows:

RESULT 2. *There is a contrast effect (as for H3) when game matrices are incomplete and the practice stage game harmony is low.*

RESULT 3. *Controlling for Result 2, cooperation is increasing in the practice stage game harmony because of an assimilation effect (as for H2).*

RESULT 4. *Result 3 is significant only because of the Guess condition: this suggests that this assimilation may be stronger when game matrices are incomplete, although the evidence for H4 is otherwise unfavourable*.

5.3 Test of H5: Forming Payoff Guesses

Do subjects make payoff guesses entirely at random? A question in the final questionnaire asked subjects to state how much they agreed with a statement that they had picked up numbers at a random, on a scale between 0 (minimum) and 10 (maximum): the mean valuation was just 2.660 (this was not significantly different across treatments). This contrasts with other answers in the final questionnaire, with much higher ratings: for example, a question asking how clear the instructions were received a mean rating of 8.266.

We can compare the histogram of mean random guesses by subject given in Treatment 8 (n = 60) to the mean guesses of the comparable first Stage 2 round of the Guess condition (Treatments 4-6, n = 48).

(Insert Figure 4 about here).

A non-parametric Epps-Singleton test shows that the two distributions are significantly different from one another [χ^2 (4) = 15.932, $P < 0.05$]. The same result holds if we compare the mean random guesses in Treatment 8 with the mean guesses for all rounds in Treatments 4-6 (n = 480): in an Epps-Singleton test, χ^2 (4) = 33.242 ($P < 0.001$). Whatever they are doing, subjects in Treatments 4-6 are not just generating numbers at random.

If subjects do not make payoff guesses at random, we would expect that they use past and current information about the game to formulate guesses about the expected payoff values in each payoff cell. Three subjects do not: they just always chose 50.[10] If we exclude them from the sample, we can use

[10] One of these three subjects put 40 in one hidden cell out of the sixty faced between round 1 and 10.

regression analysis to explore the determinants of δ.[11] We started with a model with a large set of independent variables, and iteratively eliminated those that failed to pass a significance test[12]. Details on the set of independent variables used in the most general model are included in Appendix B. Table 4 presents the final model.

(Insert Table 4 about here).

The variables in Table 4 can be classified in four groups. First, there are some personal variables (*DAge*, *HardSciences* and *AmbiguityAv*). *HardSciences* is a dummy equal to 1 for students in the hard sciences. *AmbiguityAv* is an index of ambiguity aversion built using the final questionnaire (see Appendix B): the larger *AmbiguityAv*, the smaller the measured ambiguity aversion.

Second, there are variables based purely on game features: *MeanNRD* and *SdNRD* respectively refer to the mean and standard deviation of payoffs not on the right diagonal. The coefficient on *SdNRD* is interpretable as an assimilation effect: a larger standard deviation of available payoffs can be interpreted as a larger variability of right-diagonal payoffs, which may produce a larger DI value.

Third, there are variables based purely on game antecedents: *St1Payoff*, *St1NegDiff* and *High*. *St1Payoff* is the average payoff obtained by the subject in Stage 1.[13] *St1NegDiff* is the average difference between own own's and the coplayer's gains in Stage 1, if negative (otherwise it is equal to zero). The coefficients on these variables suggest that having earned more fictional money in Stage 1 made the subject more likely to perceive things as entailing scope for cooperation, but lagging behind worked as powerful negative reinforcement. *High* (*Low*) is equal to 1 in the High (Low) condition, and 0 otherwise. The coefficient on *High* is indicative of the assimilation effect, and that on Low of the contrast effect at work.

---

[11] A random-effects regression model was used to handle individual-specific effects, as each subject made ten choices and there is evidence suggesting their existence (using a Bresch-Pagan Lagrange multiplier test) and the adequacy of a random-effects specification (using a Hausman specification test).

[12] Some variables were constructed using questionnaires given at the end of the experiment. Since these questionnaires were given only from the second session, the regression models do not use the observations from the first session (and so n = 410).

[13] Note that practice stage gains were not convertible in real money, and everybody knew it, so the significance of the coefficient on St1Payoff cannot be interpreted as a wealth effect.

Fourth, there are interaction variables between experimental condition and game harmony outside the right diagonal; a main effect variable for game harmony outside the right diagonal was removed in the reduction process. *Medium* × *GHnRD* is equal to *Medium* × G(Γ) computed over the visible part of the payoff matrix; similarly for *HighGH* × *GHnRD*. While the largest coefficient is on *MediumGH* × *GHnRD*, the coefficient on *HighGH* × *GHnRD* is also significant. Not only does experience with high game harmony in the practice stage leads to a cooperative frame, interpreted as a perception of more harmonious interests in subsequent game playing, but also the effect is stronger when what is known about the game lends itself better to a cooperative interpretation.

Clearly, the data is noisy, as shown by the low within-subject R^2. One obvious source of noise is likely to be rounding: 0.253 of the guesses involved multiples of 10, over twice as many as we would expect with a uniform distribution of guesses between 0 and 100 ($P < 0.001$), while 0.433 involved multiples of 5, versus a null of 21% ($P < 0.001$). Equally clearly, though, subjects did not simply choose numbers at random. Further, δ values agree with data on cooperation and similarity rating suggesting that in the Low, Guess condition a contrast effect was at work, and a small but significant correlation can be found between δ values and cooperative choices: lower perceived payoff inequalities between oneself and the coplayer were associated to greater cooperation (Pearson $r = -0.118$, Spearman $\rho = -0.124$, $d.f. = 449$, $P < 0.01$).

RESULT 5. *As for H5, subjects do not make payoff guesses entirely at random. It is possible to explain about half of the between-subject variance.*

RESULT 6. *The practice stage game harmony affected δ 1) either by contrast (in moving from Medium to Low) or by assimilation (in moving from Medium to High); 2) in the Medium and High conditions, by making agents more sensitive to the level of game harmony as emerging from the visible part of the payoff matrix.*

RESULT 7. *As for H5, payoff guesses were related, directly and indirectly, to actual cooperation.*

5.4 Test of H6: Problem Ambiguity

It is not the case that cooperation is affected by having subjects choose guesses when the game matrix is incomplete. Treatment 5 (Medium, Guess) and Treatment 7 (Medium, Frozen) had cooperation rates of 0.400 and 0.469, respectively (see Table 1): the difference is insignificant [t (308) = 1.219, *n.s.*]. However, we found in section 5.2 that a F test using H and I as factors is significant for I. These findings lead us to conclude that having it is having incomplete payoff matrices that probably leads to more cooperation.

RESULT 8. *Cooperation increases in the ambiguity of the decision problem. This is compatible with H6.*

How does problem ambiguity affect the use of non-cooperative algorithms by the subjects? Define P(e_m) as the frequency with which a unique action prescribed by the algorithm m is identifiable in a given sample over the visible or guess-integrated payoff matrix (in the Visible and Guess conditions, respectively). Define P(m | e_m) as the (conditional) frequency that the unique m action is followed given that a unique e_λ exists, and P (m) as the (unconditional) frequency that the unique m action is followed. Table 4 displays the values of P(e_m), P(m | e_m) and P (m) for the following non-cooperative algorithms: 1) Nash; 2) minmax; 3) "0-level strict dominance" (D0); 4) "1-level strict dominance" (D1); 5) rationalizability (D2); 4) "pure sum of payoff dominance" (L1); 5) "best response to pure sum of payoff dominance"(L2); 6) "maximum payoff dominance" (MPD). 1 and 2 are well known, and 3 and 4 are simply levels of reasoning in the rationalizability process, with rationalizability equating to "2-level strict dominance" in a 3×3 game. MPD corresponds to going for the highest conceivable payoff for itself; L1 to choosing the best action against a uniformly randomizing opponent; L2 to choosing the best action against a L1 player. A fuller description of these algorithms can be found in Zizzo and Sgroi (2000). The values are displayed for the Visible and the Guess conditions only (Treatments 1-6), as non-cooperative solutions cannot be found with the incomplete and uncompleted payoff matrices of the Frozen condition.

(Insert Table 5 about here).

The interesting result from Table 5 is that, whereas by construction all Visible game matrices have a unique Nash equilibrium, P(e_{Nash}) drops by over half over the payoff-integrated game matrices, and the dominance-based algorithms also do not fare well.[14] Conversely, simpler algorithms, such as L1 and L2, remain basically stable in their P(e_m) values.

P(m | e_m) mostly tend to be lower. This could be for two reasons. First, even with guesses perfectly corresponding to the expected payoff values, in making decisions subjects may rely not just on the expected values but also on higher moments of their subjective distribution of possible payoff values in each cell. Second, subjects may simply tremble more in the presence of problem ambiguity.

P (m) values are widely different in ways that reflect the differential success of the more complex and the simpler algorithms in prescribing a unique action in the Guess condition. Nash can only explain about one fourth of the choices, whereas algorithms such as MPD, Maxmin, L2 and especially L1 hover at about or above the 50% success rate.

RESULT 9. *As subjects face more uncertain decision problems, their behavior can be better described by simpler, less cognitively demanding strategies, as for H6, as long as they are not based on strict dominance.*

These results appear roughly compatible with the experimental findings of Costa-Gomes et al. (2001) and the computer simulations of Zizzo and Sgroi (2000), which suggest that L1 and L2 are the strategies that best correspond to what people do when facing new normal-form games: indeed, we find that this is truer when subjects face incompletely determined games than otherwise.

6. Conclusions

This paper presented an experimental exercise in cognitive game theory: we employed techniques, such as payoff integration and similarity evaluations, that could be used in further research to open the black box of framing effects. This is worthwhile since framing effects are poorly understood by economists, and yet they can have a significant impact on cooperation. We analyzed the effect of one specific and minimal framing effect: the manipulation of the harmony of the games faced in the practice stage. We analyzed this framing effect as affecting the way the game is perceived, and thus within a more general framework of a study of how game perception affects behavior. Zizzo

[14] The drops are statistically significant at $P < 0.001$.

(2002) provides a theoretical example of how this framework can be formalized, in relation to the concept of game harmony with normal form games. We used 3 × 3 games, as these better allow to discriminate among non-cooperative algorithms.

We found a strong association between game harmony and cooperation, if weaker than the one identified by Zizzo and Tan (2002). Higher game harmony in the practice stage tends to induce more cooperative behavior later on, by *assimilation* to the harmony of the previous games, but we show how this assimilation effect can be moderated by a counteracting *contrast* effect. Framing effects depend upon the balance between the two. Further, the framing effect would be smaller, and insignificant, were it not for the treatments where the game is ambiguous, i.e. when the payoff matrix is presented in an incomplete form. The framing effect operates not only by changing the *level* of cooperation, but also the *sensitivity* of cooperation to the harmony of the game. Cooperation increases in the ambiguity of the decision problem, and so does the describability of behavior with simpler algorithms than Nash and rationalizability, such as the L1 and L2 that were so successful in Costa-Gomes et al. (2001). Asking subjects to guess the payoff values, or to evaluate the similarity between games, can be a way of gathering information about how the game is perceived, and on how this affects behavior.

Appendix A – Experimental Instructions

This appendix contains, in order: 1) the instructions for the Guess condition (Treatments 4-6); 2) the changes in the case of the Visible (Treatments 1-3) and Frozen conditions (Treatment 7); 3) the instructions for the guessing task in Treatment 8.

Instructions for the Guess Condition

Instructions for the Practice Stage

You are about to participate in an experiment on decision-making. The experiment will be conducted in four stages. Stage 1 (the Practice Stage) is for practice only, while in the Payment Stage you are paid whatever amount you have earned in Stages 2 and 3, plus additional 4 pounds for participation.

In the Practice Stage you will be asked to choose *actions* for six rounds. Each round your action will be paired with that of one other participant (your *coparticipant*), and this will determine the outcomes both for you and your coparticipant. The nature of the decision in the Practice Stage is discussed below. You will always be matched with the same coparticipant in the Practice Stage. After each round you will be told what actions were chosen by you and your coparticipant, and how many *experimental points* you and your coparticipant earned in the round as the result of your actions. You will receive *no information* about the actions of and points earned by the participants that are not your coparticipant, and similarly they will not be informed about your actions or the points you have earned. In the later stages of the experiment, you will *not* be matched with the same coparticipant as in the Practice Stage.

You should try to make the best decisions you can in the Practice Stage: by doing so you can get the greatest understanding on how to do well in the rest of the experiment.

The Decision Table in the Practice Stage

Each decision that you face will be described by a *Decision Table* consisting of eighteen numbers arranged in three rows and three columns. Decision Tables will appear also in Stage 2 and Stage 3, and so it is quite important that you get a good understanding of what they represent.

An example (namely, the Decision Table for round 1) is currently on display on the computer screen. You and your coparticipant have three available actions, A, B and C. A yellow and a blue cell, in pairs, are placed in a grid in correspondence of each of the nine combinations of possible actions by you and your coparticipant. Two numbers, one in the yellow cell and one in the blue cell, are placed in correspondence of each combination of possible actions. The number in the *yellow* cell is the amount of experimental points that *you* would get for each combination of possible actions; the number in the *blue* cell is the amount of experimental points that your *coparticipant* would get for each combination of possible actions. To make some examples based on the Decision Table on the computer display: if you choose A and your coparticipant chooses C, you get [*number of points*] points and your coparticipant gets [*number of points*]; if you choose B and your coparticipant chooses B, you get [*number of points*] points and your coparticipant gets [*number of points*]; finally, if you choose C and your coparticipant chooses A, you get [*number of points*] points and your coparticipant gets [*number of points*]. The *point numbers* in all cells are always between [*number of points*] and [*number of points*].

To choose an action, you need to click one of the buttons labelled A, B and C. You should click A if you want to choose action A, B if you want to choose action B, and C if you want to choose action C. A message window will then appear asking you to confirm your choice. To do so, click OK on the window and then click the Confirm button. If you want to cancel your choice, click OK on the window and then click the Cancel button.

You will not get to know the choice of your coparticipant for the round until your coparticipant has chosen as well, and similarly he or she will not learn about your action until he or she has made his or her choice. In making your choices, you are not allowed to speak to other participants or communicate in any other way.

Before starting the practice, we ask you to answer a brief questionnaire, with the only purpose of checking whether you have understood the instructions. Raise your hand when you have completed the questionnaire.

Many thanks for your participation to the experiment, and good luck!

Please raise your hand if you have any questions.

Instructions for Stage 2

In Stage 2, you are asked to choose actions for ten rounds in relation to Decision Tables, exactly as in the Practice Stage. However, you are not told what the point numbers are in relation to some combinations of actions: that is why you find red question marks rather than numbers in some cells. Let us call these cells with question marks as the *hidden cells*, meaning that the point numbers are hidden. While you are not told the exact point numbers in hidden cells, we can tell you that they are always between 0 and 100. In all rounds, the hidden cells will be those in correspondence of the following combinations of actions: you choosing A and your coparticipant choosing C; both you and your coparticipant choosing B; you choosing C and your coparticipant choosing A. Decision Tables will change from round to round, exactly as in the Practice Stage.

At the start of the Stage 2, you will be matched with a *different* coparticipant from the one you will have played the Practice Stage with. You will have to take decisions for ten rounds, but you will receive no feedback on their outcome until the end of the experiment.

This is the last interactive stage of the experiment: your Stage 3 earnings will depend only on your choices, not on combinations of choices by you and some other participant, while Stage 4 is just for payment.

Your choices

You are asked to make your choices in two stages.

First, you should replace the question marks with the point numbers that you guess correspond to each hidden cell: let us label these point numbers as your *guesses*. To do this, click once on each hidden cell with the mouse and then use the keyboard to cancel the question mark and replace it with your best estimate of the correct value. Each guess has to be between 0 and 100. Numbers have to be typed in their arabic (1, 2, 3...) rather than verbal (one, two, three...) form. If you make some mistake and want to reset any hidden cell, just double click on it with the mouse.

Second, after having filled all the hidden cells with perceived point numbers, you can choose an action exactly as you did in the Practice Stage, first by clicking on the A, B or C button and then by confirming. You will not be allowed to make your choice of action unless you have replaced the question marks with perceived point numbers, following the instructions above. You and your coparticipant will earn point numbers as the result of your actions, exactly as in the Practice Stage.

You will not receive any feedback about the outcome of your choices after each round. No communication of any kind with the other participants is allowed.

Your winnings

Your Stage 2 winnings come from two sources: the *guess payment* (between 0 and 3 pounds) and the *action payment* (between 0 and 5 pounds). The guess payment and the action payment are determined on the basis of different rounds, randomly chosen by the computer. These payments rounds will be the same for you and your coparticipant.

Guess payment. The computer picks up your guess in a single randomly determined hidden cell and compares it with the correct point number. If you get the guess exactly right, you earn **3 pounds**. The more incorrect is your guess, the less you gain: in particular, for every point by which your guess is incorrect, you lose **twelve pence**. If your guess is wrong by 25 points or more, your guess payment is zero. The following table tells you what the guess payment is for various levels of error:

Guess Payment Table

Error (=gap between guess and correct point number)	Guess Payment (in pounds)
0	3
1	2.88
2	2.76
3	2.64
4	2.52
5	2.4
6	2.28
7	2.16
8	2.04
9	1.92
10	1.8
11	1.68
12	1.56
13	1.44
14	1.32
15	1.2
16	1.08
17	0.96
18	0.84
19	0.72
20	0.6
21	0.48
22	0.36
23	0.24
24	0.12
25 or more	0

Example: assume that, in relation to the single hidden cell randomly picked by the computer to determine your guess payment, your guess is 80 but the correct point number is 60. Then the error (i.e., the gap between your guess and the correct point number) is 20, and your guess payment is equal to 60 pence.

Note that your guesses have *no effect* on the correct point numbers corresponding to the game: hence, putting a high number, say 100, in yellow cells would **not** by itself be a good way to earn more points (unless you actually believed that a point number of 100 was hidden in that cell).

You should try to fill the hidden cells with the numbers that you actually expect to be present in each cell. This is in your own best interest, because by doing so: a) you are more likely to earn a higher guess payment, and b) it will then be easier for you to choose the action that you believe best.

Action payment. The computer will randomly choose a second payment round to determine the action payment. This second payment round will be different from the one used to determine the guess payment.

The action payment depends on the point numbers you earn in this second payment round, and so it depends on the actions by you and your coparticipant. More specifically, *each point earned in this round is worth 0.05 pounds in the Payment Stage* (so, for example, 100 points are worth 5 pounds).

Please do not take decisions in a hurry: you can improve your chances to do well by thinking carefully about each Decision Table.

Before starting making decisions, we ask you to answer a second brief questionnaire, once again with the only purpose of checking whether you have understood the instructions. Raise your hand when you have completed the questionnaire.

Please raise your hand if you have any questions.

Instructions for Stage 3

In Stage 3 you are asked to ***evaluate how similar two Decision Tables are***. The enclosed sheet contains a printout of the Comparison Decision Table (CDT). On the CDT, for each combination of actions by you and your coparticipant, the left numbers are your points, whereas the right numbers are your coparticipant's points. For example, if you were asked to take an action on the CDT, and both you and your coparticipant chose B, you would get 15 and your coparticipant would get 25. However, you are not asked to choose actions in this stage. Rather, you are asked to compare the CDT to the Decision Table that, round by round, appears on the computer screen.

Stage 3 has 16 rounds. Each round you should assess the similarity of the Decision Table on the computer screen to the CDT you have on paper: you should evaluate similarity on a scale between 0 (very different) to 100 (very similar). The similarity payment, discussed below, will depend on how accurate your similarity evaluation is, and can be up to 6 pounds.

Once you decide your similarity evaluation, you should:

- click on the white cell using the mouse;

- write down your similarity evaluation, which must be between 0 and 100, and must be typed in its arabic (1, 2, 3...) rather than verbal (one, two, three...) form;
- click on the Confirm button;
- click OK on the message box that will then appear;
- if you are satisfied with your choice, click on the Confirm button again without changing the number. Otherwise, you can click Cancel or change the number.

If you make some mistake and want to reset the white cell, just double click on it with the mouse. Any choice you make will not be communicated to the other participants, and similarly you will not learn anything about their choices.

The Similarity Payment

At the end of Stage 3, the computer will randomly choose a round and compare your choice for that round with the correct answer. If you get the evaluation exactly right, you earn **6 pounds**. The more incorrect is your evaluation, the less you gain: in particular, for every point by which your guess is incorrect, you lose **twenty-four pence**. If your evaluation is wrong by 25 points or more, your similarity payment is zero. The table on the following page tells you what your similarity payment is for various levels of error [*table was on following page*]:

Similarity Payment Table

Error (=gap between your valuation and correct answer)	Similarity Payment (in pounds)
0	6
1	5.76
2	5.52
3	5.28
4	5.04
5	4.8
6	4.56
7	4.32
8	4.08
9	3.84
10	3.6
11	3.36
12	3.12
13	2.88
14	2.64
15	2.4
16	2.16
17	1.92
18	1.68

19	**1.44**
20	**1.2**
21	**0.96**
22	**0.72**
23	**0.48**
24	**0.24**
25 or more	**0**

Example: assume that your similarity evaluation is 80 but the correct answer is 60. Then the error (i.e., the gap between your similarity evaluation and the correct answer) is 20, and your similarity payment is equal to 1.20 pounds.

It is in your own best interest to choose a similarity evaluation as accurate as possible, because by doing so you are more likely to earn a higher similarity payment.

Before starting making decisions, we ask you to answer another brief questionnaire, with the only purpose of checking whether you have understood the instructions. Raise your hand when you have completed the questionnaire.

Please raise your hand if you have any questions.

<u>Instructions for the Visible Condition</u>

[*The Stage 1 and Stage 3 instructions were identical to those of the Guess condition. The Stage 2 instructions are reproduced below.*]

Instructions for Stage 2

In Stage 2, you are asked to choose actions for ten rounds in relation to Decision Tables, exactly as in the Practice Stage. However, you are not told what the point numbers are in relation to some combinations of actions: that is why you find red question marks rather than numbers in some cells. Let us call these cells with question marks as the *hidden cells*, meaning that the point numbers are hidden. While you are not told the exact point numbers in hidden cells, we can tell you that they are always between 0 and 100. In all rounds, the hidden cells will be those in correspondence of the following combinations of actions: you choosing A and your coparticipant choosing C; both you and your coparticipant choosing B; you choosing C and your coparticipant choosing A. Decision Tables will change from round to round, exactly as in the Practice Stage.

At the start of the Stage 2, you will be matched with a *different* coparticipant from the one you will have played the Practice Stage with. You will have to take decisions for ten rounds, but you will receive no feedback on their outcome until the end of the experiment.

This is the last interactive stage of the experiment: your Stage 3 earnings will depend only on your choices, not on combinations of choices by you and some other participant, while Stage 4 is just for payment.

Your choices

You can choose an action exactly as you have done in the Practice Stage, first by clicking on the A, B or C button and then by confirming. You and your coparticipant will earn point numbers as the result of your actions, exactly as in the Practice Stage.

You will not receive any feedback about the outcome of your choices after each round. No communication of any kind with the other participants is allowed.

Your winnings

The computer will randomly choose a payment round to determine the *action payment*. This payment round will be the same for you and your coparticipant.

The action payment depends on the point numbers you earn in this round, and so it depends on the actions by you and your coparticipant. More specifically, *each point earned in this round is worth 0.05 pounds in the Payment Stage* (so, for example, 100 points are worth 5 pounds).

Please do not take decisions in a hurry: you can improve your chances to do well by thinking carefully about each Decision Table.

Before starting making decisions, we ask you to answer a second brief questionnaire, once again with the only purpose of checking whether you have understood the instructions. Raise your hand when you have completed the questionnaire.

Please raise your hand if you have any questions.

Instructions for the Frozen Condition

The Stage 1 and Stage 3 instructions were identical to those of the Guess and Visible conditions. The first section of the Stage 2 instructions were identical to the corresponding one for the Guess condition; the sections under the headings "Your choices" and "Your actions" were identical to the corresponding ones for the Visible condition.

Instructions for the Guessing Task in Treatment 8

In the guessing task, you have to choose six numbers between 0 and 100 (you can put the same number more than once, if so you wish). You choose numbers by clicking once on each hidden cell with the mouse and then using the keyboard to cancel the question mark and replace it with your choice. Numbers have to be typed in their arabic (1, 2, 3...) rather than verbal (one, two, three...) form. If you make some mistake and want to reset any hidden cell, just double click on it with the mouse. After having filled all the white cells with your guesses, you can finalise your decisions by clicking the Numbers Chosen button and then confirm by clicking it again.

The computer then picks up your guess in a single randomly determined hidden cell and compares it with the winning (correct) guess. If you get the guess exactly right, you earn *3 pounds*. The more incorrect is your guess, the less you gain: in particular, for every point by which your guess is incorrect, you lose *twelve pence*. If your guess is wrong by 25 points or more, your guess payment is zero. A table enclosed to these instructions tells you what the guess payment is for various levels of error. [*The table was identical to the equivalent one for the Stage 2 instructions in the Guess condition, as reproduced above*.] You will not learn how much you won in the guessing task until the end of the experiment.

Appendix B – Variables used in most general model

i) Personal Variables

DAge: age minus mean age.
Sex: 1 for males, 0 for females.
Economics, *Humanities*, *HardSciences*: 1 for subjects with an Economics, humanities, hard sciences background (respectively), 0 otherwise.
QAverage: final questionnaire evaluation by subject on their agreement (from 0 to 10) with the statement about the payoff guess task "I was not sure about the number, so I chose the likely average each time".
QRandom: same as *QAverage* in relation to statement "I just picked numbers at random".
QSimilarity: same as *QAverage* in relation to statement: "I chose numbers on the basis of the similarity of the Decision Table to previous Decision Tables".
RiskAv: risk aversion index. The final questionnaire contained five hypothetical binary choices between A) 50 pounds per choice and B) 100 pounds with probability k% and 0 pounds with probability $(1 - k)$%, for k = 50, 60, 70, 80 and 90, which we can respectively map into the index K = 1, 2, 3, 4 and 5. Let the answer to the question concerning K be $a(K)$, and let $a(K) = 1$ if B (the risk-loving choice) is chosen. Then $RiskAv = K \times a(K)$: the lower the *RiskAv*, the more risk averse the agent is likely to be.
AmbiguityAv: ambiguity aversion index. The final questionnaire contained five hypothetical binary Ellsberg choices where the subject wins if she picks a green ball out of one of two bags with 100 balls of one of two colors, green and blue. She does not know how many green or blue balls there are in Bag 2, but she knows that Bag 1 has k green balls and $(1 - k)$ blue balls, for k = 35, 40, 45, 50, 55, which we can respectively map into the index K = 1, 2, 3, 4 and 5. Let the answer to the question concerning K be $a(K)$, and let $a(K) = 1$ if Bag 2 (the ambiguity-loving choice) is chosen. Then $AmbiguityAv = K \times a(K)$: the lower the *AmbiguityAv*, the more ambiguity averse the agent is likely to be.

ii) Game-Specific Variables

MeanNRD: mean of the visible payoffs (i.e. those not on the right diagonal).
SdNRD: standard deviation of the visible payoffs (i.e. those not on the right diagonal).

iii) Game Antecedents Variables

High: 1 in Hight condition, 0 otherwise.
Low: 1 in Low condition, 0 otherwise.
Round: Stage 2 round a game is played in.
St1Payoff: mean payoff obtained by the subject in Stage 1.
St1PayoffOther: mean payoff obtained by the subject's coplayer in Stage 1.
St1NegDiff: mean difference between own and coplayer's payoff in Stage 1, if negative; otherwise 0.

iv) Interaction Variables

Low × *GHnRD*: *Low* × G(Γ) computed over the visible payoff matrix
Medium × *GHnRD*: *Medium* × G(Γ) computed over the visible payoff matrix
High × *GHnRD*: *High* × G(Γ) computed over the visible payoff matrix

References

Buschena, D., Zilberman, D., 1999. Testing the effects of similarity on risky choice: implications for violations of expected utility. Theory and Decision 46, 251-276.

Cookson, R., 2000. Framing effects in public good experiments. Experimental Economics 3, 55-79.

Costa-Gomes, M., Crawford, V.P., Broseta, B., 2001. Cognition and behavior in normal-form games: an experimental study. Econometrica 69, 1193-1235.

Croson, R.T.A., 2000. Thinking like a game theorist: factors affecting the frequency of equilibrium play. Journal of Economic Behavior and Organization 41, 299-314.

Elliott, C.S., Hayward, D.M., Canon, S., 1998. Institutional framing: some experimental evidence. Journal of Economic Behavior and Organization 35, 455-464.

Gilboa, I., Schmeidler, D., 2001. A theory of case-based decisions. Cambridge University Press, Cambridge.

Hahn, U., Chater, N., 1997. Concepts and similarity. In: K. Lamberts and D. Shanks (Eds.), Knowledge, Concepts and Categories. Psychology Press, Hove, pp. 43-92.

Kahneman, D., Varey, C., 1991. Notes on the psychology of utility. In: J. Elster and J. Roemer (Eds.), Interpersonal Comparisons of Well-Being. Cambridge University Press, Cambridge, pp. 127-159.

Mitropoulos, A., 2001. Learning under minimal information: an experiment on mutual fate control. Journal of Economic Psychology 22, 523-557.

Morris, M.W., Sim, D.L.H., Girotto, V., 1998. Distinguishing sources of cooperation in the one-round Prisoner's Dilemma: evidence for cooperative decisions based on the illusion of control. Journal of Experimental Social Psychology 34, 494-512.

Oechssler, J., Schipper, B., 2000. Can you guess the game you're playing? Mimeo, University of Bonn.

Pruitt, D.G., Kimmel, M.J., 1977. Twenty years of experimental gaming: critique, synthesis, and suggestions for the future. Annual Review of Psychology 28, 363-392.

Zizzo, D.J., 2002. On the measurement of harmony in normal-form games. Department of Economics Discussion Paper 116, University of Oxford.

Zizzo, D.J., Sgroi, D., 2000. Bounded-rational behavior by neural networks in normal form games. Nuffield College Oxford Economics Discussion Paper No. 2000-W30.

Zizzo, D.J., Tan, J.H.W., 2002. Game harmony as a predictor of cooperation in 2×2 games. Department of Economics Discussion Paper 117, University of Oxford.

Fig. 1. A computer display from the Guess and Frozen conditions in Stage 2.

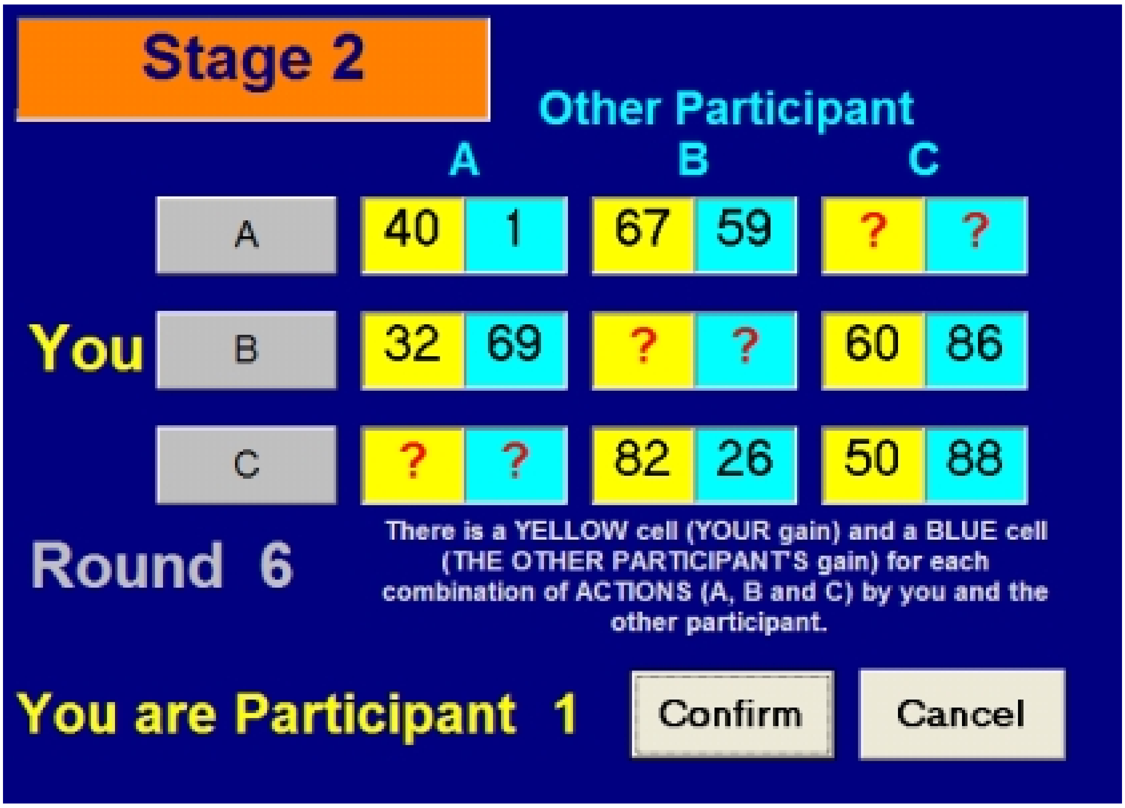
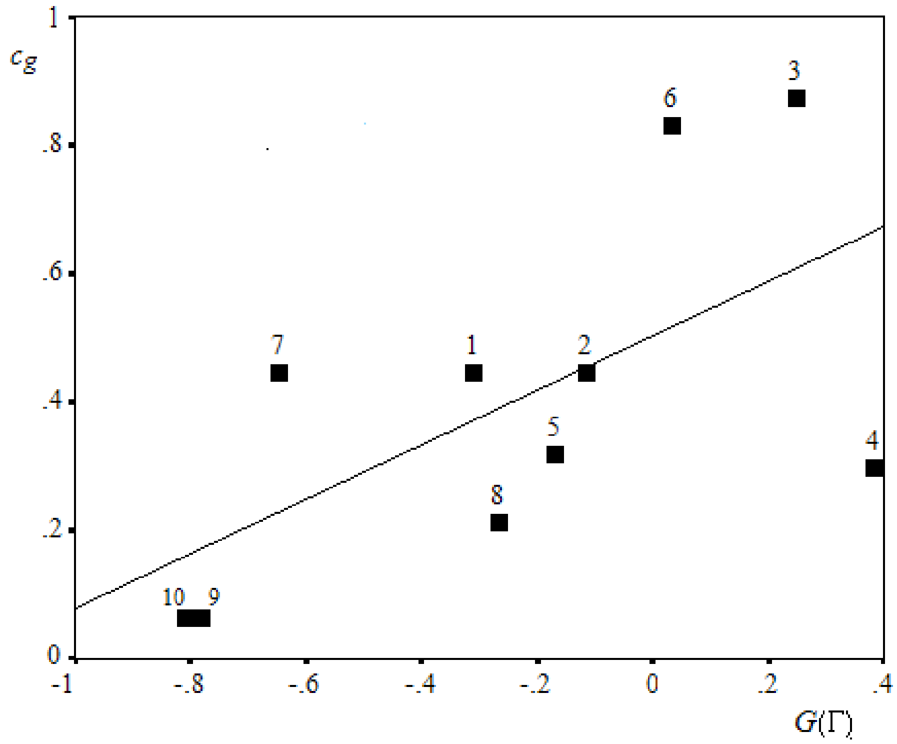
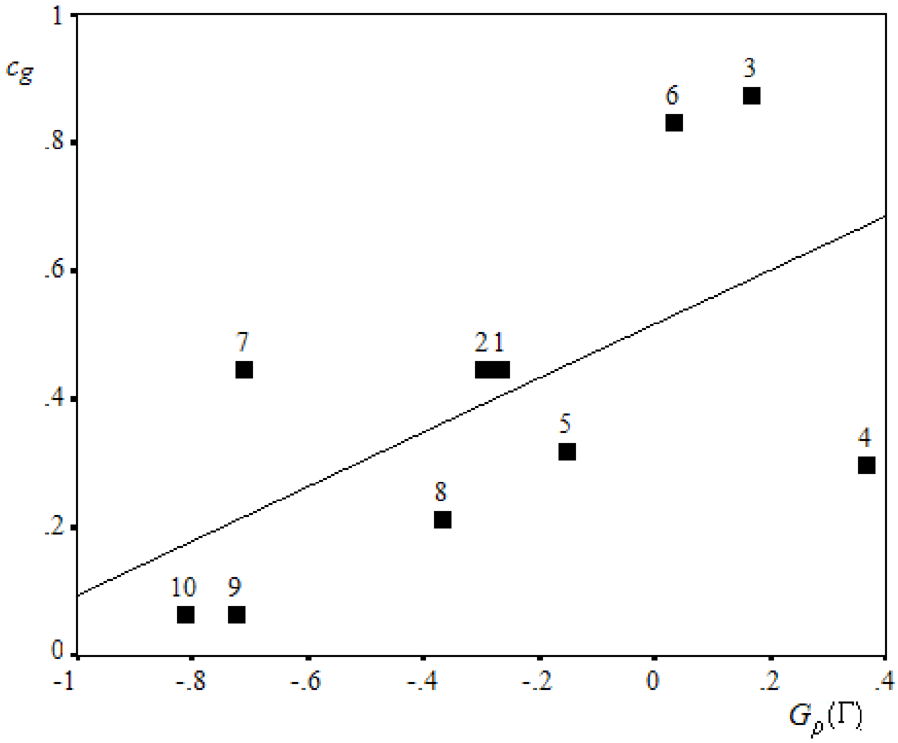


Fig. 2. The Comparison Decision Table as presented on paper to the subjects.

		Your coparticipant A		B		C	
You	A	5	56	88	52	72	43
	B	36	61	15	25	57	41
	C	86	22	23	23	77	28

Fig. 3. Scatterplots of game harmony on cooperation rate by game.[a]

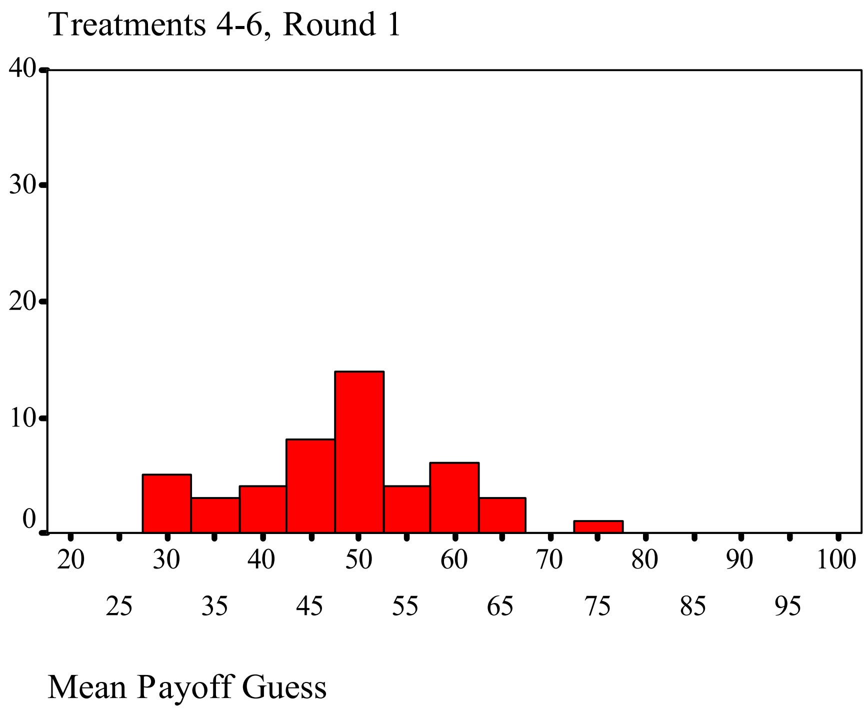


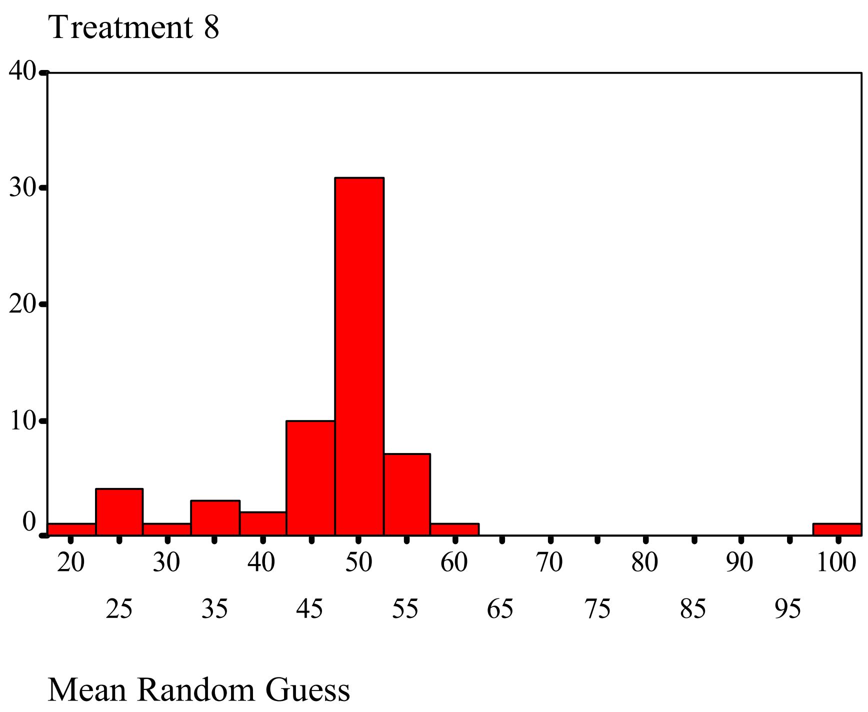


[a] The observation labels correspond to the games, as specified by the round of presentation in Table 2.

Fig. 4. Histograms of mean guesses by subject.

Fig. 4. Histograms of mean guesses by subject.[a]





[a] The vertical axys has the number of observations for every 5 guess points interval.

Table 1
Experimental treatments [a]

Treatment	*H* Condition	*I* Condition	n	c_t	δ	s_t
1	Low	Visible	16	0.363	-	39.895
2	Medium	Visible	16	0.406	-	40.109
3	High	Visible	15	0.433	-	35.396
4	Low	Guess	16	0.475	32.989	31.596
5	Medium	Guess	16	0.469	35.123	37.816
6	High	Guess	16	0.581	26.034	42.445
7	Medium	Frozen	15	0.4	-	38.342
8	-	-	61	-	-	-

[a] *H* condition refers to whether the treatment had High, Medium or Low game harmony in the practice stage. *I* condition refers to whether the payoff matrix was fully displayed (Visible), incomplete (Frozen), or incomplete and with a request to guess payoff values (Guess). c_t is the mean cooperation rate by treatment. δ is the mean of a summary statistic of the difference between guess of own's own and the coplayer's payoff. s_t is the mean evaluation of similarity between games in Stage 3, by treatment.

Table 2.
Stage 1 and 2 games [a]

	Column Player		
Row	A, B	C, D	E, F
Pl.	G, H	I, J	K, L
	M, N	O, P	Q, R

	Round	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
Practice	1	4	87	47	39	19	53	98	8	3	63	42	67	25	53	69	21	38	62
Low	2	22	78	26	64	11	96	63	1	83	18	12	72	16	55	83	1	17	82
Game	3	33	91	72	31	80	16	28	62	52	56	81	7	12	100	32	45	58	12
Harmony	4	36	43	16	77	93	6	43	85	35	56	68	38	16	66	87	17	86	3
Games	5	19	68	46	31	86	15	23	78	43	54	56	24	31	72	29	65	40	59
	6	20	35	96	8	30	40	93	13	90	8	82	1	9	93	10	74	28	42
Practice	1	72	12	11	3	83	95	26	36	94	69	27	1	45	100	36	87	90	27
Medium	2	23	52	18	80	28	74	96	98	98	26	28	46	56	76	26	0	78	59
Game	3	84	80	70	10	23	71	22	54	3	21	50	15	69	97	6	59	63	94
Harmony	4	66	66	72	39	33	19	22	30	18	3	59	95	83	73	66	7	93	92
Games	5	24	68	25	69	16	10	72	19	93	57	42	94	24	21	51	24	84	26
	6	31	92	93	17	51	20	61	4	11	68	69	84	97	69	11	93	2	57
Practice	1	86	96	59	41	44	5	21	4	59	60	24	6	65	85	35	7	42	11
High	2	96	86	4	21	85	65	41	59	60	59	7	35	5	44	6	24	11	42
Game	3	15	19	64	77	88	87	99	96	8	14	95	84	18	65	2	11	11	23
Harmony	4	23	31	96	84	25	20	31	10	91	75	8	24	13	4	61	80	10	12
Games	5	95	95	55	43	22	2	99	60	44	7	71	56	33	8	12	12	63	38
	6	77	82	57	67	27	33	63	85	30	29	64	42	90	97	20	17	89	89
Stage 2	1	60	3	79	11	*58*	*7*	20	47	*44*	*52*	68	93	*6*	*71*	27	59	98	48
Games	2	26	39	30	5	*24*	*99*	57	12	*28*	*12*	90	59	*51*	*47*	10	72	36	55
	3	76	71	83	29	*45*	*3*	37	84	*15*	*53*	29	51	*93*	*82*	42	2	71	68
	4	59	54	40	9	*18*	*85*	0	29	*79*	*71*	88	57	*58*	*88*	20	29	3	43
	5	70	18	24	89	*67*	*75*	5	65	*67*	*62*	7	10	*92*	*5*	7	66	89	75
	6	40	1	67	59	*70*	*77*	32	69	*44*	*88*	60	86	*45*	*11*	82	26	50	88
	7	54	61	72	55	*78*	*67*	78	30	*68*	*3*	31	96	*36*	*71*	48	71	49	100
	8	21	16	14	76	*44*	*10*	21	47	*31*	*69*	42	65	*40*	*88*	22	89	48	30
	9	40	40	61	31	*81*	*21*	31	58	*48*	*52*	70	40	*20*	*80*	40	69	61	62
	10	69	69	39	86	*12*	*100*	83	39	*55*	*56*	25	69	*100*	*10*	69	24	40	41

[a] Games are defined row-by-row: to obtain the payoff matrix corresponding to a given round and condition, replace the A, B, C, D… values in the generic payoff matrix with the corresponding value on the row for that round and condition. For the Stage 2 games, the (right diagonal) payoff cells that were hidden in Treatments 4-7 are displayed in italics.

Table 3.
Stage 2 games [a]

Round 1		
60, 3	79, 11	58, 7
20, 47	44, 52	68, 93
6, 71	27, 59	98, 48

Round 6		
40, 1	67, 59	70, 77
32, 69	44, 88	60, 86
45, 11	82, 26	50, 88

Round 2		
26, 39	30, 5	24, 99
57, 12	28, 12	90, 59
51, 47	10, 72	36, 55

Round 7		
54, 61	72, 55	78, 67
78, 30	68, 3	31, 96
36, 71	48, 71	49, 100

Round 3		
76, 71	83, 29	45, 3
37, 84	15, 53	29, 51
93, 82	42, 2	71, 68

Round 8		
21, 16	14, 76	44, 10
21, 47	31, 69	42, 65
40, 88	22, 89	48, 30

Round 4		
59, 54	40, 9	18, 85
0, 29	79, 71	88, 57
58, 88	20, 29	3, 43

Round 9		
40, 40	61, 31	81, 21
31, 58	48, 52	70, 40
20, 80	40, 69	61, 62

Round 5		
70, 18	24, 89	67, 75
5, 65	67, 62	7, 10
92, 5	7, 66	89, 75

Round 10		
69, 69	39, 86	12, 100
83, 39	55, 56	25, 69
100, 10	69, 24	40, 41

[a] In the Visible and Frozen conditions, the right diagonal on each matrix was hidden by question marks (see Figure 1 for an example).

Table 4.
Random effects regression on δ

	Coef.	SE	Sign.
DAge	0.56	0.18	***
HardSciences	5.805	2.868	**
AmbiguityAv	-7.696	3.15	**
SdNRD	0.325	0.136	**
High	-11.167	4.396	**
Low	-12.065	5.458	**
St1Payoff	-0.283	0.139	**
St1PayoffOther	-0.242	0.137	*
St1NegDiff	0.373	0.159	**
Medium x GHnRD	-11.859	3.311	***
High x GHnRD	-6.943	3.087	**
Constant	76.189	15.311	***
Within R^2	0.045		
Between R^2	0.552		
Overall R^2	0.267		

Table 5.[a]
Performance of non-cooperative algorithms

	Visible Condition			Guess Condition		
Algorithm	$P(e_m)$	$P(m \mid e_m)$	$P(m)$	$P(e_m)$	$P(m \mid e_m)$	$P(m)$
Nash	1	0.702	0.702	0.463	0.581	0.269
Minmax	1	0.668	0.668	0.929	0.567	0.527
D0	0.353	0.855	0.302	0.098	0.936	0.092
D1	0.502	0.869	0.436	0.154	0.77	0.119
D2	0.651	0.85	0.553	0.223	0.738	0.165
L1	0.951	0.745	0.709	0.971	0.631	0.613
L2	0.949	0.684	0.649	0.969	0.536	0.519
MPD	0.9	0.66	0.594	0.973	0.563	0.548

[a] $P(e_m)$ as the rate with which a unique action prescribed by the algorithm m is identifiable. $P(m \mid e_m)$ is the rate that the unique m action is followed given that a unique e_λ exists. $P(m)$ is the rate that the unique m action is followed.