DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

GAME HARMONY AS A PREDICTOR OF COOPERATION IN 2 × 2 GAMES

Daniel John Zizzo

Jonathan H. W. Tan

Number 117

September 2002

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

Manor Road Building, Oxford OX1 3UQ

Game Harmony as a Predictor of Cooperation in 2 × 2 Games

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

Jonathan H.W. Tan
Lincoln College
University of Oxford

(Version September 3, 2002)

Abstract

This paper presents an experimental test of the relationship between *game harmony* and cooperation in 2 × 2 games. Game harmony measures describe how harmonious (non-conflictual) or disharmonious (conflictual) the interests of players are, as embodied in the payoffs: coordination games and constant-sum games are examples of games of perfect harmony and disharmony, respectively, with most games being somewhere in the middle. In our experiment we consider a variety of 2 × 2 games, including amongst others the Prisoner's Dilemma, the Stag-Hunt, the Chicken, a coordination game and three variants of trust games. We find that simple game harmony measures can explain 2/3 (or more) of the variance in mean cooperation rate across these games.

Keywords: game harmony, cooperation, 2 × 2 games, trust games, social dilemmas.
JEL Classification Codes: C72, C91, H41.

* Department of Economics, University of Oxford, Manor Road Building, Manor Road, Oxford OX1 3UQ, U.K. (email: daniel.zizzo@economics.ox.ac.uk; jonathan.tan@economics.ox.ac.uk). Please address all correspondence to the first author. We thank Michael Bacharach and David Buschena for advice and Jahirul Islam for exprimental assistance. The usual disclaimer applies. This paper is dedicated to the memory of Michael Bacharach.

This paper presents an experimental test of the relationship between *game harmony* and cooperation in 2×2 games, including familiar ones such as the Prisoner's Dilemma, the Stag-Hunt, the Chicken and three variants of trust games. Game harmony is a generic game property that describes how harmonious (non-conflictual) or disharmonious (conflictual) the interests of players are, as embodied in the payoffs. Consider the coordination game, constant-sum game and Prisoner's Dilemma of Table 1.

(Insert Table 1 about here).

In the coordination game there is perfect harmony of interests between the players: the only problem is one, indeed, of coordination. In the constant-sum game, the gain of a player is the loss of another, which means that there is perfect *dis*harmony of interests. Most games, however, do not fall in either of these two extremes: rather, they are somewhere in the middle, in terms of harmony of interests, between coordination games and constant-sum games. Game harmony measures are just algorithms that make this "somewhere in the middle" precise. Zizzo's (2002) most basic measure is an average of Pearson correlation coefficients among payoff pairs: in the case of a 2×2 game, it reduces to the simple Pearson correlation coefficient between the payoffs of the players for each state of the world. Another measure replaces Pearson with Spearman correlations, and so a cardinal with an ordinal measure of game harmony. The ordinal measure is less sensitive than the cardinal one to the particular payoff values being chosen, and so might be better when one seeks to compare game harmony between different game structures; for robustness, we consider both measures. In two-players games, both of them are bounded between -1 and $+1$, with $+1$ the value for coordination games and -1 the value for constant-sum games: higher numerical values indicate greater game harmony.

Our experiment tested three hypotheses on how game harmony affects behavior in a variety of 2×2 games. We list them in increasing order of empirical support found. First, we hypothesized that, if a subject plays a set of games of particularly high or low game harmony, this may distort behavior in later games being played. No support was found for this hypothesis.

Second, we predicted that, the more similar a subject perceives a game to a perfectly harmonious relative to a perfectly disharmonious game benchmark, the more cooperatively she will play: this is because she may perceive the game as entailing more harmony of interests. Our measure of similarity is a proxy for the degree of perception of harmony of the game. We found some support for this similarity hypothesis.

Our third hypothesis was perhaps the most basic: the greater the harmony of a game, the more the cooperation we should expect in a game. There is strong support for this hypothesis in our experimental data: either of our game harmony measures can explain some two thirds of the variance of the mean cooperation rate (or more, depending on how the data is aggregated). While game harmony is clearly not the only thing that matters for cooperation, we believe that it has a significant role to play. We are not aware of other attempts trying to summarize some dimension of "cooperativeness" in a numerical index in a general setting, although Anatol Rapaport and Albert Chammah (1965) devised specific cooperation indices in relation to 2×2 Prisoner's Dilemmas.

Cooperation is important in a variety of economic settings, such as social dilemmas, the functioning of businesses and organizations, and any transaction where economic agents have to trust and fulfill trust in order for the Pareto optimal outcome to be achieved. We believe it is important to identify properties of games that are easy to estimate and easy to use as predictors of cooperation, such as measures of game harmony. Furthermore, in economies with different levels of social capital there may be a tendency to perceive games with the same material payoff structure as entailing different levels of perceived harmony and hence cooperation: our technique based on similarity judgments can be a way of probing into this and improving our understanding of social capital.

In section 1 we review the main relevant results of Zizzo (2002) and present our experimental hypotheses. Section 2 describes the experiment, section 3 presents the results and section 4 concludes. Appendices A and B contain the experimental instructions and core raw data, respectively.

I. Experimental Background

A. *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 $W = W_1 \times \mathrm{W}_2 \times ... \times W_n$ 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 $w \in W$. Under some weak technical assumptions guaranteeing existence (see Zizzo, 2002), one 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:

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

In the case of 2 × 2 games, $C = 1$, so G(Γ) = $r(a_w, b_w)$, which is obviously bounded between -1 and +1. We shall mention three of the properties that can be proven in relation to G(Γ). G(Γ) is scale-independent, meaning that G(Γ) = G(x_i, W) = G($k_i\ x_i + h_i$, W) for any finite $k_i > 0$ and h_i. G(Γ) 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). Further, 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:

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

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.

PROPOSITION (Equivalence Result). *Assume that the material payoffs a_{cw} and b_{cw} are given. Then* $G(\Gamma^v) \gtreqless G(\Gamma^u) \leftrightarrow \beta_i \gtreqless 0\ \forall\, i$, *and* $\partial\beta_i / \partial G(\Gamma^v) > 0$.

The proof of this result is in Zizzo (2002). It 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 to attain it (such as in the Prisoner's Dilemma). We do not believe, however, that this is the only way in which a relationship between cooperation and game harmony can be established. This paper tries a different approach, experimental rather than theoretical: our experiment verifies whether game harmony can indeed be used as a predictor of cooperation in games, and in what sense this is the case.

A closely related measure of game harmony can be obtained by considering payoff orderings rather than their 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:

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

In the case of 2 × 2 games, this reduces to $G_{\rho}(\Gamma) = r_c(a_{cw}^{\rho}, b_{cw}^{\rho})$. Both G(Γ) and $G_{\rho}(\Gamma)$ implicitly assign equal weight to all states of the world. A more general class of game harmony measures in Zizzo (2002) relaxes this assumption: in order for these more general measures to have any bite, however, one would need to specify how the different weights are determined. A natural way of doing this would be to assume that, if a subject follows an algorithm $\mathcal{L}$ (e.g., Nash) in solving Γ, a weight of 0 should be assigned to the $\mathcal{L}$-dominated outcomes, and game harmony should be defined only over the $\mathcal{L}$-undominated outcomes. While this may have some normative appeal, it also leads to a "uniqueness paradox": if $\mathcal{L}$ has a unique solution, then there is only one $\mathcal{L}$-undominated outcome, and so game harmony cannot be computed, since no correlation coefficients among payoffs can be computed on the basis of just one observation. This is paradoxical, as it means that, the more successful $\mathcal{L}$ is in pinpointing a unique solution, the less the range of games over which $\mathcal{L}$-weighted game harmony can be computed. This makes $\mathcal{L}$-weighted game harmony unusable in our context, since many well-known 2 × 2 games have unique solutions with algorithms such as L1 or L2 (as defined by Dale Stahl and Paul Wilson, 1995, and Miguel Costa-Gomes et al., 2001) or Nash, and even a weak algorithm such as rationalizability would imply the non computability of game harmony in the Prisoner's Dilemma.

In their general applicability and lack of degrees of freedom, we believe that G(Γ) and $G_{\rho}(\Gamma)$ have the virtues of simplicity and parsimony. It will be on them that we shall focus in the rest of our paper to make a case for game harmony as a predictor of cooperation in 2 × 2 games.

B. *Experimental Hypotheses*

A claim that game harmony matters for cooperation is a claim that the harmony of the game, *as perceived* by the subject, has an impact on her behavior. But there is no reason to expect that the game perceived by a subject is the same as the material payoff matrix assigned to her by the experimenter: for example, what the experimenter defines as a Prisoner's Dilemma might actually be a rather different game, say a Stag-Hunt, in the eyes of the subject.

Let us assume, nevertheless, that there is strong link between perceived harmony and game harmony as can be estimated from the game matrix. If so, one can estimate $G(\Gamma)$ and $G_\rho(\Gamma)$ from the material payoff matrix, treating the material payoffs as the utility values x_{iw}, and use $G(\Gamma)$ and $G_\rho(\Gamma)$ as predictors of cooperation in Γ:

H1. Whenever a unique cooperative action is defined in Γ*, the mean cooperation rate (i.e., the mean probability of choosing this action) is increasing in* $G(\Gamma)$ *and* $G_\rho(\Gamma)$.

The beauty of H1 is its informational simplicity: we need no information apart from that provided by the 2×2 payoff matrix in order to have some idea of what cooperation rate to expect. However, this may not necessarily be the case. Subjects may perceive the game as being more or less harmonious than what its material payoff structure entails, depending on situational factors such as the way the instructions have been phrased or a history of play of similar games, and more individual-specific factors such as the player's personality. *Perceived* game harmony, insofar as it is related to the psychologically transformed payoff matrix and this does not coincide with the material payoff matrix, is unobservable. Let us define $T(\Gamma)$ and $T_\rho(\Gamma)$ as the "true" game harmony values, i.e. those corresponding to the $G(\Gamma)$ and $G_\rho(\Gamma)$ (respectively) of the utility payoffs matrix however this is perceived by the subject. We need some additional experimental tool, apart from asking subjects to play in Γ, to gather information on how changes in purely perceived game harmony, i.e. in $\Delta T(\Gamma)$ or $\Delta T_\rho(\Gamma)$, affect cooperation.

In stage 3 of our experiment, subjects had to rate incentive-compatibly how similar each stage 2 game was to either a perfectly harmonious game (a coordination game) or a perfectly disharmonious game (a constant-sum game). A *similarity index S* can be built to measure the degree to which Γ is assimilated to the perfect harmony relative to the perfect disharmony benchmark (see section 3): changes in *S* for a given payoff matrix Γ can be used as a proxy for $\Delta T(\Gamma)$ or $\Delta T_\rho(\Gamma)$. If so, then an increase in $\Delta T(\Gamma)$ or $\Delta T_\rho(\Gamma)$ should result both in a rise in *S* and in a rise in cooperation.

H2. For a given payoff matrix Γ, *the cooperation rate is increasing in* $T(\Gamma)$ *and* $T_\rho(\Gamma)$, *as proxied by S. More specifically*:

1. *higher S values should be correlated with a higher likelihood of cooperation for any given game and subject (strong similarity hypothesis)*;
2. *subjects with higher average S are subjects that cooperate more on average (weak similarity hypothesis)*.

There are two versions of H2. The strong one is compatible with the idea that subjects actually assimilate each game to a more cooperative or less cooperative benchmark in order to decide what choice to take, and so to a causal rather than simply correlational link between *S* and cooperation. It may be thought akin to case-based decision reasoning (Itzhak Gilboa and David Schmeidler, 2001) as applied to game theory and combined with the idea of game harmony as a relevant dimension of similarity.

The weak similarity hypothesis simply states that people who, for whatever idiosyncratic reason such as a personality trait, tend to perceive games more harmoniously tend also to cooperate more.

We tested whether distortions in $T(\Gamma)$ and $T_\rho(\Gamma)$ are of meaningful size also in another way, leading to H3. We varied the harmony $G(\Gamma)$ and $G_\rho(\Gamma)$ of the practice stage (Stage 1) games, while retaining the same set of Stage 2 (and Stage 3) games: this was to see whether this simple procedural manipulation had a significant effect on later perception of game harmony and thus on cooperative behavior.

H3. The more the $G(\Gamma)$ *and* $G_\rho(\Gamma)$ *of practice stage games, the greater* $\Delta T(\Gamma)$ *or* $\Delta T_\rho(\Gamma)$ *(as proxied by S) and so the greater the mean cooperation rate.*

This can be thought of as a kind of framing manipulation, entailing an impact of a previous task on a later task. There are contexts where this is known to occur. For example, if subjects first read some news briefs on cooperatives business strategies and have to answer a few questions on this topic, they contribute more in the follow-up play of a public good contribution game (Catherine Elliott et al., 1998).

II. The Experimental Design

Fifteen experimental sessions were run in the Department of Economics in the University of Oxford in the March of 2002. Each session had four subjects, for a total of an overall 60 subjects. The experiment was divided in three stages (plus the payment stage), all of which involved 2×2 games. The games were never exactly symmetric, though they were sometimes approximately so (as will be explained below), so one could play each game in one of two roles. It may be tempting to label these roles as that of "row player" and that of "column player", but subjects always had game matrices displayed on the screen in such a way that they would be row players: this was achieved by suitably transposing game matrices as needed. We thus find more appropriate to define the role according to whether, in any given round, subjects saw the "direct" (standard) or "transposed" presentation of the game on the computer display, and label their roles as that of d-players and t-players, respectively. All game payoffs were provided as numbers between 0 and 100.

In the experimental instructions games were labeled as "decision tables", players as "participants" and coplayers as "coparticipants". 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. Instructions and questionnaires were provided on paper, but the experiment was otherwise computerized.

Stage 1. In the first stage ("Practice Stage"), subjects did practice by playing six 2×2 games twice, namely once as d-players and once as t-players. They were matched with a single co-player throughout the stage, played games in random order and received feedback about the outcome of each round. Practice stage points did not count towards final winnings. Games had been chosen according to the following procedure: 1) payoff values were generated randomly, by drawing payoff values from a uniform distribution between 0 and 100; 2) games without a unique pure Nash equilibrium were discarded; 3) according to the experimental condition, games with high, medium or low game harmony were selected (High, Medium and Low condition, respectively). Games with a unique pure Nash equilibrium were chosen because we wanted games to be, at least roughly, of the same strategic complexity notwithstanding differences in game harmony. Furthermore, in order to ensure that any distortion in behavior in the later stages were not due to a particular reinforcement learning history from the practice stage, we tried to use different game samples in different sessions, albeit games of about the same level of game harmony for sessions of the same condition.[1] The mean $G(\Gamma)$ [$G_\rho(\Gamma)$] values were 0.985 (0.943), 0.020 (0.013) and -0.980 (-0.915) for the games used in the High, Medium and Low conditions, respectively. Each of the two practice stage pairs in each session played the games in random order.

Stage 2. In the second stage, subjects played ten 2×2 games twice, again once as d-players and once as t-players. Games were played in random order. No feedback was received after each round in this stage. At the end of the experiment, a round was randomly picked by the computer to determine the "action payment". The subject's action in this round was matched with that of a different "coparticipant" from that of the practice stage. Each payoff point earned in the payment round was converted into 0.06 U.K. pounds. Subjects could therefore earn between zero and six pounds as their action payment, depending on their choice of action, that of their coplayer and the game played in the payment round. The game matrices are displayed in Table 2.

(Insert Table 2 about here).

Most Table 2 games are payoff-perturbed versions of familiar 2×2 games: Prisoner's Dilemma (PD), Stag-Hunt (StH) and Chicken (ChK), plus a constant-sum game (CSG) and a coordination game (CDG). The Envy (Altruism) Game (EG and AG, respectively) is a game with a strictly dominant solution in its material payoff values: deviations from the strictly dominant solution can be interpreted as due (if not to trembling) to envy (altruism) or other negative (positive) interdependence in preferences. The remaining games are instances of trust games as defined in

[1] Originally 4 sessions were planned for each condition and so four sets of games were prepared for each condition. When a fifth session was run at about the same time because of a surplus of subjects, one of the sets of games was used again for each condition. Thus, in each condition one game sample was used twice. Twelve different game samples were used in a total of fifteen sessions.

James Coleman (1990) and Michael Bacharach et al. (2001): the d-player is the truster, with a choice whether to trust (by playing top) or withholds trust (by playing bottom); the t-player is the trustee, with a choice whether to fulfill or violate the trust. Bacharach et al. (2001) employ both the Kind Trust Game (KTG) and the Needy Trust Game (NTG): the difference between the two is that, in the latter, the truster is “more in need”, i.e. she ends up in a worse fate if she does not trust. UTG stands for Unequitable Trust Game and is equivalent to the KTG, but with about 20 points added to the payoff values of the truster and 30 points subtracted from the payoff values of the trustee.

Small payoff perturbations have been used before in experiments with 2 $\times$ 2 games (e.g. Frederick Rankin et al., 2000). They were used in our experiment for two reasons: 1) we wanted to reduce the likelihood that subjects would realize that they were playing each game in both roles; 2) we wanted games in stage 2 not to appear “different” from Stage 1 games because of the symmetric nature of some of them and the higher frequency of the same numbers being used as payoff values. Payoff perturbation explains why games were never exactly symmetric, although four out of ten (PD, StH, ChK, CDG) were basically so (i.e., were symmetrical up to the payoff perturbation).

Stage 3. In the third stage, subjects had to evaluate how similar a game was to one of two other games, the so-called “comparison decision tables” (CDT). The CDTs were the constant-sum game and coordination game from Stage 2, as for Table 2. The CDTs were compared with the other eight games of Table 2, both in their “direct” and in their transposed form, plus the transpose of the CDTs, plus another constant-sum game and a coordination game (see Table 3).

(Insert Table 3 about here).

Thus, each of the two CDTs was compared with twenty games, and there were an overall forty rounds with as many similarity evaluations. The order of presentation was randomized.

Following David Buschena and David Zilberman (1999), similarity evaluations were given on a Likert scale between 0 and 9, with 0 indicating maximum similarity and 9 maximum dissimilarity (i.e., *lower* values on the scale pointed to *higher* similarity). The Stage 3 “Similarity Payment” was determined on the basis of the similarity evaluation of a randomly chosen round as follows. Subjects were paid £ 12 to get the evaluation exactly exactly right, with a penalty of one £ 5 per each point of error (subjects were paid zero for getting the guess wrong by three points or more). Thus, a standard absolute difference incentive-compatibility mechanism was implemented (e.g., Rachel Croson, 2000). The instructions contained a table with details on the amount of payment for any given level of error, so subjects were not required to make any significant computation; also, a couple of questions in the practice questionnaire were used to check their understanding of the mechanism. Payment was determined at the end of the experiment and subjects received no feedback on the outcome of their similarity choices during Stage 3. For this reason, the determination of the "correct" similarity answer was an issue that had to be practically addressed to ensure financial motivation and determine payments, but not one with serious bearing on the experiment. This was important in the light of the arbitrariness of any formula for "correct" similarity values, and the potential distortions that learning feedback could have produced.[2]

Payment Stage. At the end of the experiment subjects were paid the action payment (up to U.K. £ 6), the similarity payment (up to £ 12), plus £ 4 for participation. Average payments were about £ 10 for around one hour of work.

III. Experimental Results

A. *Definitions*

In order to determine whether game harmony was associated with a higher likelihood of play of the cooperative outcome, we need to define what the cooperative outcome is for each game in Stage 2. In the light of the games at hand, this is not particularly difficult: unique cooperative actions, when defined, are illustrated in bold in Figure 1. There is no cooperative action defined in the constant-sum game, nor (ignoring the payoff perturbation) there is a clear one for the t-player in the EG and AG (although payoff perturbation implies a strictly dominant solution in both cases). For the d-player, the altruistic action in the AG and the non-envious action in the EG are considered as the cooperative actions; in the case of the AG, this can be justified, for example, on the basis of either a utilitarian social welfare function or the Nash bargaining solution. For the trust games, the (*trust*, *fulfilment*) pair is always identified as the cooperative outcome. In the coordination game the payoff (and Pareto) dominant solution is taken as the cooperative outcome. The cooperative

[2] 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 and a CDT. Then the "correct" answer was determined as the integer number closest by rounding to $4.5 \times Unique + 4.5 \times [(EDP - min\,(EDP)] / [max(EDP) - min(EDP)]$.

outcomes are obvious in the remaining cases. There are an overall 960 Stage 2 observations where a cooperative outcome is defined.

In order to test H2 and H3, we need to identify a *similarity index S* measuring the degree to which Γ is assimilated to the perfect harmony relative to the perfect disharmony benchmark. Call s_{CDG} the similarity evaluation of a game relative to the coordination game CDG, and s_{CSG} the similarity evaluation relative to the constant-sum game CSG (in both cases, lower values entail greater similarity). It is then natural to define a similarity index $S = s_{CSG} / (s_{CDG} + s_{CSG})$, bounded between 0 and 1, with higher values indicating closer similarity to the perfectly harmonious relative to the perfectly disharmonious game benchmark.

B. *Results*

Evidence for H3. We find no support for the hypothesis that differences in practice stage experience due to the game harmony manipulation had any significant effect on behavior. Mean cooperation rates were 0.469, 0.550 and 0.525 in the Low, Medium and High condition, respectively: the differences are not only non-monotonic, but also not significant [F (2, 957) = 2.222, $P > 0.1$]. Mirroring this result, there is no significant difference in mean S across conditions: it is equal to 0.528, 0.512 and 0.522 in Low, Medium and High, respectively [F (2, 957) = 0.950, $P > 0.1$]. We can conclude that there is no support for H3 in our dataset of 2×2 games.

Evidence for H1. Table 4 presents the the mean cooperation rate for each game and the corresponding game harmony values. There is a striking correlation between game harmony, both in its cardinal and its ordinal measure, and mean cooperation rate c_Γ for each game ($P < 0.001$,[3] $n = 9$): Pearson r[c_Γ, G(Γ)] = 0.921, r[c_Γ, $G_\rho(\Gamma)$] = 0.923, Spearman ρ[c_Γ, G(Γ)] = 0.962, ρ[c_Γ, $G_\rho(\Gamma)$] = 0.911. These correlation coefficients imply R^2 above 0.8 in (ordinary or rank) regressions of game harmony values on mean cooperation rates. Figure 1 illustrates the strong correlation for both G(Γ) and G_ρ (Γ), and, at the same time, some of its limitations.

(Insert Table 4 and Figure 1 about here).

The main outliers in the case of G(Γ) are the KTG and UTG: while these games are only just slightly more disharmonious than the Chicken game, subjects are only about half as likely to cooperate in them. As an example of "anomaly" in the case of $G_\rho(\Gamma)$, Chicken, Stag-Hunt and NTG

[3] Here and below, all tests of the significance of correlation coefficients are two-tailed.

have all about the same cooperation rate even though they differ in $G_\rho(\Gamma)$ values, from lowest to highest respectively. But of course these are anomalies only if we were to believe what cannot be believed, namely that game harmony is the only thing that matters in determining cooperation. For example, the strategic structure of the Chicken game is clearly different from that of asymmetric trust games such as the KTG, NTG and UTG, and this may not be fully captured by the game harmony values.

A general limitation of the analysis so far is that unweighted game harmony cannot explain differences in cooperation rates between different roles in asymmetric games: by grouping cooperation rates by game and so averaging out cooperation rates in different roles in the asymmetric games, we are ignoring this source of unexplained variation. Further, the AG and EG cooperation rates are computed from half of the observations that we have for the other games of Figure 1, since they are computed on the basis of d-players' choices only (the only ones for which, as we explained above, we can define a cooperative choice for these games). This is not taken into account in the analysis above.

We can address both of these limitations by computing cooperation rates by game role, rather than just by game, as illustrated by Table 4. Figure 2 draws the corresponding scatterplots of game harmony values on cooperation.

(Insert Figure 2 about here).

While, as we we would expect, correlation coefficients are lower, they are still above 0.8 ($P < 0.001$, $n = 16$): Pearson $r[c_r, G(\Gamma)] = 0.817$, $r[c_r, G_\rho(\Gamma)] = 0.834$, Spearman $\rho[c_r, G(\Gamma)] = 0.864$, $\rho[c_r, G_\rho(\Gamma)] = 0.838$. The main loss of explanatory power seems to come from the NTG, the only game with a wide difference in cooperation rates between roles (trusters almost always trust, but trustees fulfill trust only three times out of ten). With this qualification, our unweighted game harmony measures still retain considerable explanatory power: a simple OLS regression of $G(\Gamma)$ on c_r gives $\beta = 0.350$ (S.E. = 0.066, $R^2 = 0.668$), and one of $G_\rho(\Gamma)$ on c_r gives $\beta = 0.362$ (S.E. = 0.064, $R^2 = 0.695$). Thus, we find strong support for H1. In our dataset of 2×2 games, some two thirds of the variance in mean cooperation rates can be explained by a single, simple predictor such as an unweighted game harmony measure: a 0.1 increase in game harmony should increase the cooperation rate by some 3.5%, and a unit increase by as much as 35%.

Evidence for H2. There is some evidence for the strong similarity hypothesis, although the effect does not appear quantitatively large.[4] If we define c as a dummy variable equal to 1 if a subject cooperated and to 0 otherwise, we find that Pearson $r(c, S) = 0.084$ ($P < 0.02$, $n = 840$) and Spearman $\rho(c, S) = 0.088$ ($P = 0.01$, $n = 840$). One problem in evaluating this result is that many observations may be non-independent as they come all from the same subject. Another problem is that an alternative explanation of this result is in terms of H1 rather than H2: S may just be acting as an instrument for $G(\Gamma)$ and $G_\rho(\Gamma)$, rather than genuinely capturing changes in perceived game harmony $\Delta T(\Gamma)$ or $\Delta T_\rho(\Gamma)$ and the correlation of these with cooperative outcomes.[5]

To sidestep these problems, we ran probit regressions of the form $c = \Phi(\beta' x)$, where $\Phi(y)$ denotes the probability that a standard normal variate is less than y, and where standard errors are corrected for the existence of individual-specific effects. Table 5 reproduces the results.

(Insert Table 5 about here).

As independent variables we used not only S but also a game harmony measure [$G_\rho(\Gamma)$ in Table 5, although this is inessential], a Condition variable dummy (= 0 for Low, 1 for Medium and 2 for High), a Round variable equal to the round number the observation refers to (to test for sequencing or timing effects) and a SameAsNE dummy equal to 1 if the cooperative action coincides with the unique pure Nash equilibrium action.

There are only three significant variables in Table 5 (apart from the constant in Model 1). One is SameAsNE, although with a somewhat surprising negative sign (implying less cooperation when the cooperative action coincided with Nash). Another one is $G_\rho(\Gamma)$, with a positive coefficient close to unity [this would not change if $G(\Gamma)$ were used instead]. The third variable is S, suggesting that there is a genuine positive impact of the perceived degree of harmony on cooperation, even controlling for individual-specific effects and for $G_\rho(\Gamma)$ [the use of $G(\Gamma)$ would not change this result as well].

---

[4] The mean S value was 0.521, with S.D. = 0.148.

[5] While non-significant with only 8 observations (games) for which $G(\Gamma)$ and $G_\rho(\Gamma)$ are defined, the correlations values between mean S by game and $G(\Gamma)$ [$G_\rho(\Gamma)$] are sufficiently high as to suggest this as a possible scenario: the Pearson r is 0.162 (0.102), while the Spearman ρ is 0.252 (0.180).

There is quantitatively stronger evidence for the weak similarity hypothesis. Let c_s and S_s be the mean cooperation rate and the mean S index by a subject. Subjects who tend to perceive games as more harmonious, and hence tend to assimilate games to the perfectly harmonious CDT, are more likely to cooperate on average: Pearson $r(c_s, S_s) = 0.265$ ($P < 0.05$, $n = 60$) and Spearman $\rho(c_s, S_s) = 0.313$ ($P < 0.02$, $n = 60$). Figure 3 illustrates this correlation. It cannot be explained by assuming that S_s is working as a proxy for changes in $G(\Gamma)$ and $G_\rho(\Gamma)$ rather than in $\Delta T(\Gamma)$ or $\Delta T_\rho(\Gamma)$, since each subject faces the same sample of games. Put it differently, because of the experimental design there is a zero correlation between S_s and $G(\Gamma)$ or $G_\rho(\Gamma)$.

(Insert Figure 3 about here).

IV. Discussion and Conclusions

Simple game harmony measures can be used to predict the occurrence of cooperation across a range of well-known 2×2 games, such as the Prisoner's Dilemma, the Stag-Hunt, the Chicken and three variants of trust games. Knowledge of the material payoff matrix is all that is required to compute these measures. Even when due attention is given to the meaningful asymmetry of roles in some of these games, we find Pearson and Spearman correlations above 0.8 between our measures of game harmony and mean cooperation rates.

Of course, the material payoff matrix may not be the same as the utility payoff matrix as this is perceived by economic agents. We proxied distortions in perceived game harmony relative to the material payoff matrix by asking subjects to rate how similar each game was to a coordination game (a perfectly harmonious benchmark) or to a constant-sum game (a perfectly disharmonious benchmark). We found that subjects who tend to perceive games more harmoniously are more likely to engage in cooperative behavior.

Equally obviously, it would be naïve to think that game harmony is the only thing that matters for cooperation, for at least two reasons. First, because of the scale invariance property, it does not take into account how much is at stake in cooperating when everything else is proportionately the same (e.g., Bacharach, in press). Second, two games can have the same $G(\Gamma)$ and $G_\rho(\Gamma)$ and yet be likely to lead to different cooperation rates, because of differences in the strategic structure of the game. Table 6 contains two pairs of games exemplifying these limitations.

(Insert Table 6 about here).

We still believe, on the basis of our results, that simple game harmony measures can be powerful tools in assessing cooperation rates in 2 × 2 games, including well-known and important formulations of social dilemmas and trust games.

Appendix A

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 twelve 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 eight numbers arranged in two rows and two 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 two available actions, A and B. 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 A, 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 B 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 0 and 100.

To choose an action, you need to click one of the buttons labelled A and B. You should click A if you want to choose action A, and B if you want to choose action B. 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.

<u>Instructions for Stage 2</u>

In Stage 2, you are asked to choose actions for twenty rounds in relation to Decision Tables.

At the start of 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 twenty 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 or B 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 0.06 pounds in the Payment Stage* (so, for example, 100 points are worth 6 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.

<u>Instructions for Stage 3</u>

In Stage 3 you are asked to ***evaluate how similar two Decision Tables are***. The screen displays two decision tables, a regular decision table at the top of the screen and a second decision table on the bottom of the screen. This second decision table is the Comparison Decision Table (CDT), and it is in monochrome. *For the CDT as for any decision table, 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*. 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 forty rounds. Each round you should assess the similarity of the Decision Table on the upper part of the computer display to the CDT you have on its bottom part: **you should evaluate similarity on a scale between 0 (virtually identical) to 9 (extremely different)**. Hence, the *more* similar you believe the two Decision Tables to be, the *lower* the similarity value you should assign. The similarity payment, discussed below, will depend the accuracy of your similarity evaluation, and can be up to 12 pounds.

Once you decide your similarity evaluation, you should:

- click on the white cell using the mouse (this action can only be made after 10sec into the round);
- write down your similarity evaluation, which must be between 0 and 9, 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 **12 pounds**. The more incorrect is your evaluation, the less you gain: in particular, for every point by which your guess is incorrect, you lose **5 pounds**. If your evaluation is wrong by 3 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:

Similarity Payment Table

Error (=gap between your valuation and correct answer)	**Similarity Payment (in pounds)**
0	**12**
1	**7**
2	**2**
3 or more	**0**

Example: assume that your similarity evaluation is 8 but the correct answer is 6. Then the error (i.e., the gap between your similarity evaluation and the correct answer) is 2, and your similarity payment is equal to 2 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.

Appendix B

This appendix contains the data on subjects' Stage 2 choices classified by session (Sess.), condition (Cond.: H, M and L stand for High, Medium and Low, respectively), player ID (Pl.) and game (presented in the direct and transposed form). It will be recalled that subjects were always row players of the 2 x 2 games: 1 stands for choice for the top action, 0 stands for choice for the bottom action (over the direct matrix as for Figure 2, or its transpose). S_s (mean similarity index) values, based on Stage 3 choices, are also provided (to 3 decimal places). CSG is the constant-sum game; for the other games codes, see the notes to Figure 1.

			Direct Presentation of the Game									
Sess.	Cond.	Pl.	PD	CDG	EG	AG	CSG	KTG	UTG	NTG	St-H	ChK
1	M	1	1	0	1	1	1	1	0	0	0	1
1	M	2	0	1	1	0	0	0	0	0	1	0
1	M	3	1	0	1	1	1	1	0	0	1	1
1	M	4	1	1	1	0	0	1	1	0	0	1
2	H	1	1	0	1	0	0	0	0	0	1	0
2	H	2	0	0	1	0	0	0	0	0	1	0
2	H	3	1	0	1	1	0	0	1	0	1	0
2	H	4	1	0	1	1	1	1	1	0	1	1
3	L	1	1	0	1	1	1	0	0	0	1	1
3	L	2	1	0	1	1	0	1	1	0	1	1
3	L	3	1	1	1	1	1	0	0	0	0	1
3	L	4	1	0	1	1	1	1	1	0	1	1
4	H	1	1	0	0	1	0	1	1	1	0	1
4	H	2	1	0	0	1	0	1	1	0	1	1
4	H	3	1	0	1	1	1	0	1	0	1	0
4	H	4	1	0	0	0	0	0	0	0	1	1
5	M	1	0	0	1	0	1	0	0	0	0	0
5	M	2	1	0	1	1	0	1	1	0	1	1
5	M	3	1	0	1	1	1	0	0	0	1	0
5	M	4	1	0	1	1	0	1	0	0	1	0
6	M	1	1	0	1	1	1	0	0	0	1	0
6	M	2	1	0	0	1	0	1	1	1	0	1
6	M	3	1	0	1	1	0	1	1	0	0	0
6	M	4	1	1	1	0	1	1	1	0	1	0
7	H	1	1	1	1	1	0	1	1	0	0	0
7	H	2	1	1	1	1	0	1	1	0	0	0
7	H	3	1	0	1	1	0	0	0	0	1	0
7	H	4	0	0	1	1	0	0	0	0	1	0
8	L	1	0	0	1	0	0	1	1	0	1	1
8	L	2	1	1	1	1	0	1	1	0	0	0
8	L	3	1	1	0	1	0	1	1	1	0	1
8	L	4	1	0	1	1	1	1	1	0	0	0
9	M	1	1	1	1	1	1	1	1	0	0	0
9	M	2	0	0	1	1	0	0	1	0	1	0
9	M	3	1	0	1	1	1	1	0	1	1	0
9	M	4	1	0	1	0	1	1	0	0	1	1
10	H	1	0	0	1	0	0	0	1	0	1	1
10	H	2	1	1	1	0	1	1	1	0	0	0
10	H	3	1	0	1	1	0	1	1	0	1	1
10	H	4	1	0	1	1	0	1	1	0	1	0
11	L	1	1	0	1	1	1	1	0	0	0	1
11	L	2	1	0	0	1	0	1	1	1	0	1
11	L	3	1	0	1	1	0	1	0	0	1	0
11	L	4	1	0	1	1	0	0	0	0	1	0
12	M	1	1	0	1	1	1	1	1	0	1	0
12	M	2	1	0	1	1	0	1	1	0	1	1
12	M	3	0	0	1	1	1	1	1	0	1	1
12	M	4	1	0	1	1	1	1	1	0	0	1
13	H	1	0	0	1	1	0	0	1	0	1	1
13	H	2	1	0	1	1	1	1	1	0	0	0
13	H	3	0	1	0	1	0	1	1	0	0	1
13	H	4	0	1	1	1	0	1	0	0	0	0
14	L	1	1	0	1	1	0	1	1	0	1	0
14	L	2	0	0	1	0	1	0	0	0	1	0
14	L	3	1	0	1	1	0	1	0	0	1	0
14	L	4	1	0	1	1	0	1	1	0	1	0
15	L	1	1	0	1	1	0	1	1	0	1	0
15	L	2	1	1	1	1	0	1	1	0	0	0
15	L	3	1	0	1	1	0	1	0	0	1	0
15	L	4	0	0	1	1	0	1	1	0	1	0

Sess.	Cond	Pl.	Transposed Presentation of the Game										S_s
			PD	CDG	EG	AG	CSG	KTG	UTG	NTG	St-H	ChK	
1	M	1	0	0	0	0	1	0	0	0	1	1	0.552
1	M	2	1	0	1	0	1	1	0	1	0	1	0.563
1	M	3	0	1	1	0	1	1	1	0	1	1	0.523
1	M	4	0	1	0	0	1	0	0	1	0	1	0.444
2	H	1	0	0	1	0	1	0	1	1	0	0	0.518
2	H	2	1	0	1	0	1	1	1	1	0	0	0.578
2	H	3	0	0	0	0	1	1	1	1	0	1	0.489
2	H	4	0	0	0	0	1	0	0	0	1	1	0.549
3	L	1	0	0	0	0	1	0	0	0	0	1	0.631
3	L	2	0	0	0	0	1	0	1	0	0	0	0.532
3	L	3	0	1	0	0	0	0	0	0	0	1	0.499
3	L	4	0	0	0	0	1	0	0	0	0	1	0.53
4	H	1	0	1	0	1	1	0	0	0	1	1	0.477
4	H	2	0	0	1	0	1	0	0	1	1	0	0.506
4	H	3	1	0	1	0	1	0	0	0	0	0	0.53
4	H	4	0	1	0	1	0	1	0	0	1	0	0.472
5	M	1	1	0	1	0	1	1	1	1	0	1	0.51
5	M	2	0	0	0	0	1	1	0	0	0	0	0.501
5	M	3	0	0	1	0	1	1	0	1	0	1	0.479
5	M	4	0	0	1	0	1	0	1	1	0	1	0.566
6	M	1	0	0	1	0	1	1	0	0	0	1	0.498
6	M	2	0	0	1	0	1	0	0	0	0	0	0.45
6	M	3	0	0	1	0	1	0	0	0	1	1	0.464
6	M	4	0	0	0	0	1	0	0	1	0	0	0.524
7	H	1	0	0	0	0	1	0	0	0	1	1	0.502
7	H	2	0	0	0	1	1	1	0	0	1	1	0.561
7	H	3	1	0	1	1	1	1	1	1	1	1	0.538
7	H	4	0	0	1	0	1	0	1	1	0	0	0.552
8	L	1	0	0	0	0	1	0	0	0	0	0	0.53
8	L	2	0	1	0	0	1	0	0	0	0	1	0.486
8	L	3	0	1	0	1	1	0	0	0	1	0	0.507
8	L	4	0	0	0	0	1	0	0	1	1	1	0.506
9	M	1	0	0	0	0	1	0	0	0	1	1	0.533
9	M	2	1	0	1	0	1	1	0	1	0	1	0.525
9	M	3	0	0	0	0	1	1	0	0	0	1	0.514
9	M	4	0	0	1	0	1	0	1	0	0	0	0.538
10	H	1	1	0	1	0	0	0	0	1	1	1	0.499
10	H	2	0	0	0	0	1	0	0	0	0	1	0.53
10	H	3	0	0	0	0	1	0	0	0	0	1	0.561
10	H	4	0	0	0	0	0	0	0	0	0	0	0.516
11	L	1	0	0	0	0	1	1	0	0	1	1	0.472
11	L	2	0	0	1	0	1	0	0	0	1	0	0.497
11	L	3	0	0	0	0	0	0	0	0	0	0	0.633
11	L	4	1	0	0	0	1	1	0	1	0	1	0.471
12	M	1	0	0	0	0	1	0	0	0	0	1	0.438
12	M	2	0	0	1	0	1	0	0	0	0	0	0.5
12	M	3	0	0	0	0	1	0	0	1	0	1	0.618
12	M	4	0	0	0	0	1	0	0	0	1	0	0.507
13	H	1	0	0	1	0	0	1	0	0	1	1	0.519
13	H	2	0	1	0	0	0	0	0	0	1	0	0.535
13	H	3	1	0	1	1	0	0	1	0	0	0	0.519
13	H	4	0	0	1	0	1	1	1	0	1	0	0.482
14	L	1	0	0	1	0	1	0	0	0	1	0	0.585
14	L	2	0	0	0	0	1	1	1	1	0	1	0.553
14	L	3	1	0	0	0	1	0	0	0	0	1	0.53
14	L	4	0	0	0	0	1	0	0	0	0	1	0.625
15	L	1	0	0	0	0	1	0	0	0	0	0	0.464
15	L	2	0	0	0	0	1	0	0	0	1	1	0.446
15	L	3	0	0	0	0	1	0	0	0	0	1	0.533
15	L	4	1	0	0	0	1	1	1	0	1	1	0.538

REFERENCES

Bacharach, Michael. *Beyond the individual: The theory of cognition and cooperation in games*. Princeton: Princeton University Press, in press.

Bacharach, Michael; Guerra, Gerardo and Zizzo, Daniel J. "Is Trust Self-Fulfilling? An Experimental Study." Mimeo, University of Oxford, Oxford, 2001.

Buschena, David E. and Zilberman, David. "Testing the Effects of Similarity on Risky Choice: Implications for Violations of Expected Utility." *Theory and Decision*, June 1999, *46*(3), pp. 251-76.

Coleman, James S. *Foundations of social theory*. Cambridge, MA: Harvard University Press, 1990.

Costa-Gomes, Miguel; Crawford, Vincent P. and Broseta, Bruno. "Cognition and Behavior in Normal-Form Games: An Experimental Study." *Econometrica*, September 2001, *69*(5), pp. 1193-1235.

Croson, Rachel T. A. "Thinking Like a Game Theorist: Factors Affecting the Frequency of Equilibrium Play." *Journal of Economic Behavior and Organization*, March 2000, *41*(3), pp. 299-314.

Elliott, Catherine S.; Hayward, Donald M. and Canon, Sebastian. "Institutional Framing: Some Experimental Evidence." *Journal of Economic Behavior and Organization*, May 1998, *35*(4), pp. 455-64.

Gilboa, Itzhak and and Schmeidler, David. *A theory of case-based decisions*. Cambridge: Cambridge University Press, 2001.

Rankin, Frederick W.; Van Huyck, John B. and Battalio, Raymond C. "Strategic Similarity and Emergent Conventions: Evidence from Similar Stag Hunt Games." *Games and Economic Behavior*, August 2000, *32*(2), pp. 315-37.

Rapoport, Anatol and Chammah, Albert M. *Prisoner's dilemma: A study in conflict and cooperation*. Ann Arbor: University of Michigan Press, 1965.

Stahl, Dale O. and Wilson, Paul W. "On Players' Models of Other Players: Theory and Experimental Evidence." *Games and Economic Behavior*, July 1995, *10*(1), pp. 218-54.

Zizzo, Daniel J. "On the Measurement of Harmony in Normal Form Games." Mimeo, University of Oxford, Oxford, 2002.

TABLE 1 – EXAMPLES OF GAMES WITH DIFFERENT HARMONY

Coordination Game		Constant-Sum Game		Prisoner's Dilemma	
3, 3	0, 0	0, 3	3, 0	2, 2	0, 3
0, 0	1, 1	3, 0	0, 3	3, 0	1, 1

TABLE 2 – GAMES USED IN STAGE 2 AND 3 IN THEIR "DIRECT" PRESENTATION

Prisoner's Dilemma		Envy Game		Altruism Game	
92, 11	38, 37	*61, 72*	*59, 73*	59, 28	61, 29
64, 63	10, 93	50, 28	48, 29	*48, 72*	*50, 73*

Stag-Hunt		Chicken	
10, 51	**92, 93**	92, 36	10, 11
52, 53	52, 9	**63, 62**	38, 94

Trust Games					
Kind Trust Game		Unequitable Trust Game		Needy Trust Game	
33, 34	34, 35	52, 3	53, 4	3, 34	4, 35
81, 82	14, 100	**100, 51**	33, 69	**81, 82**	14, 100

Also Comparison Decision Tables in Stage 3			
Constant-Sum Game		Coordination Game	
71, 31	18, 84	74, 75	32, 31
26, 76	89, 13	13, 12	**85, 86**

Notes: Bold letters stand for a unique cooperative solution. Italics denote the unique cooperative action for the row player in the Envy Game and the Altruism Game. Obviously in the actual experiment the cooperative actions and outcomes were not highlighted in any way. Stage 2 t-players saw the game matrices in their transposed form.

TABLE 3 – GAMES USED ONLY IN STAGE 3

Constant-Sum Game 2		Coordination Game 2	
36, 82	93, 25	61, 62	17, 16
81, 37	30, 88	4, 3	70, 71

TABLE 4 – GAME HARMONY AND MEAN COOPERATION RATES

				Direct game	Tranposed game
Game (Γ)	G(Γ)	Gρ(Γ)	c_Γ	c_r	c_r
Prisoner's Dilemma	-0.817	-0.8	0.2	0.217	0.183
Coordination Game	1	1	0.825	0.783	0.867
Envy Game	0.98	0.6	0.883	0.883	
Altruism Game	-0.98	-0.6	0.2	0.2	
Constant-Sum Game	-1	-1			
Kind Trust Game	0.066	-0.2	0.308	0.3	0.317
Unequitable Trust Game	0.066	-0.2	0.308	0.383	0.233
Needy Trust Game	0.503	0.8	0.608	0.917	0.3
Stag-Hunt	0.488	0.513	0.633	0.65	0.617
Chicken	0.16	0.2	0.6	0.583	0.617

Notes: G(Γ) and $G_\rho(\Gamma)$ are our cardinal and ordinal game harmony measures, respectively. c_Γ is the mean cooperation rate for each game, whenever a cooperative outcome is defined for at least one of the two players. Whenever defined, the "direct game" c_r is the mean cooperation rate of the row players facing the game matrices as presented in Table 2. Whenever defined, the "transposed game" c_r is the mean cooperation rate of the row players facing the transpose of the game matrices as presented in Table 2 (equivalently, they can be conceptualized as the column players of the games in Table 2).

TABLE 5 – PROBIT REGRESSION ON THE OCCURRENCE OF COOPERATION

	Model 1			Model 2		
	Coef.	Robust SE	Sign.	Coef.	Robust SE	Sign.
Condition	0.147	0.175				
S	0.629	0.302	*	0.618	0.301	*
Gρ(Γ)	0.92	0.101	**	0.913	0.1	**
SameAsNE	-0.291	0.055	**	-0.296	0.053	**
Round	0.002	0.008				
Constant	-0.508	0.244	*	-0.376	0.2	

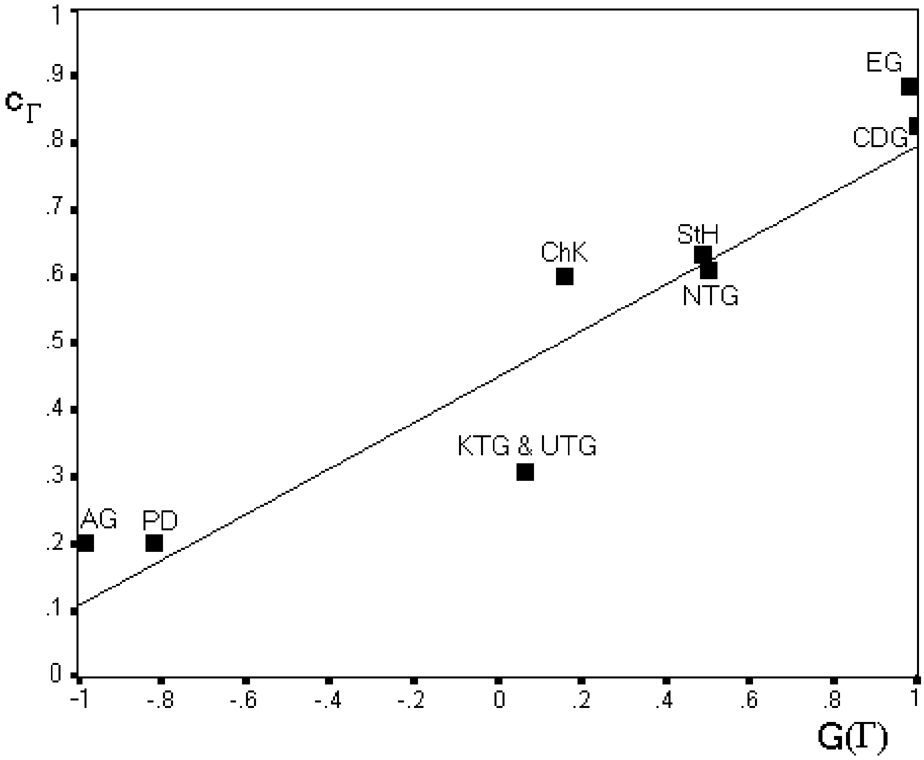
Notes: **: 0.01 significance; *: 0.05 significance. The dependent variable is $c = 1$ if a subject cooperated for any given game, and 0 otherwise; c is defined only if there is a unique cooperative action ($n = 780$). Condition = 0 in the Low, 1 in the Medium and 2 in the High condition. Round is equal to the round number the game is played in. SameAsNE = 1 if the cooperative action is the same as the unique pure Nash equilibrium action. The log-likelihood is equal to –477.198 and –478.297 for Model 1 and 2, respectively.

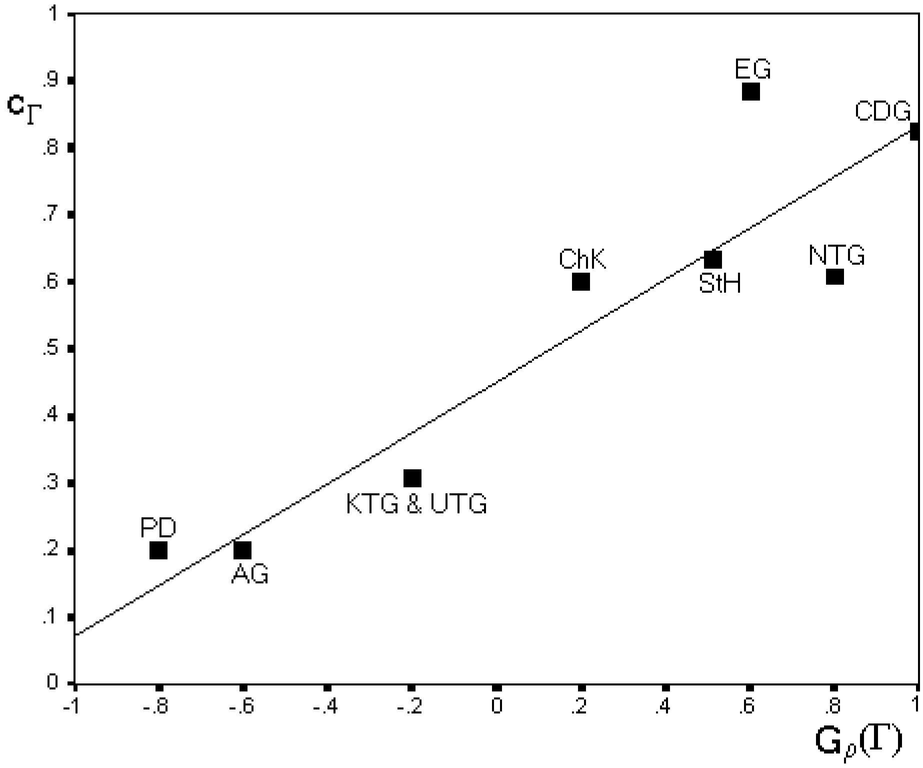
TABLE 6 – EXAMPLES OF PAIRS OF GAMES WHERE GAME HARMONY DOES NOT WORK

Example 1 (varying scope for cooperation)			
Game A		Game B	
1, 1	0, 0	100, 100	0, 0
0, 0	0.01, 0.01	0, 0	1, 1

Example 2 (changing strategic nature of game)			
Game C		Game D	
2, 2	0, 3	2, 2	0, 3
3, 0	1, 1	1, 1	3, 0

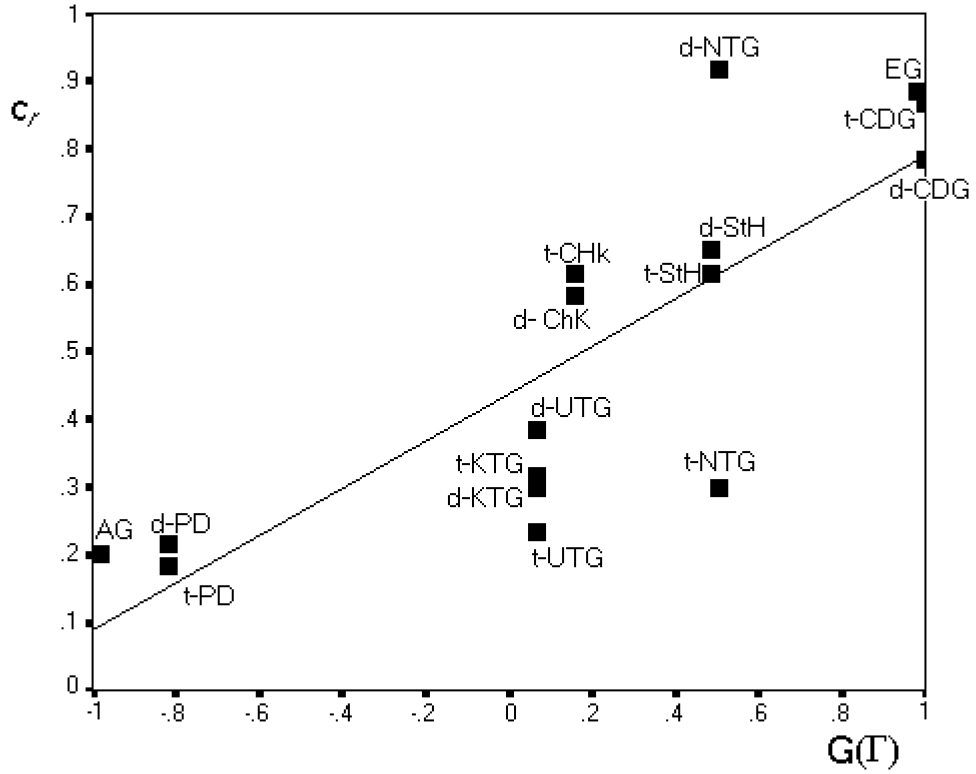
FIGURE 1. GAME HARMONY AND MEAN COOPERATION BY GAME

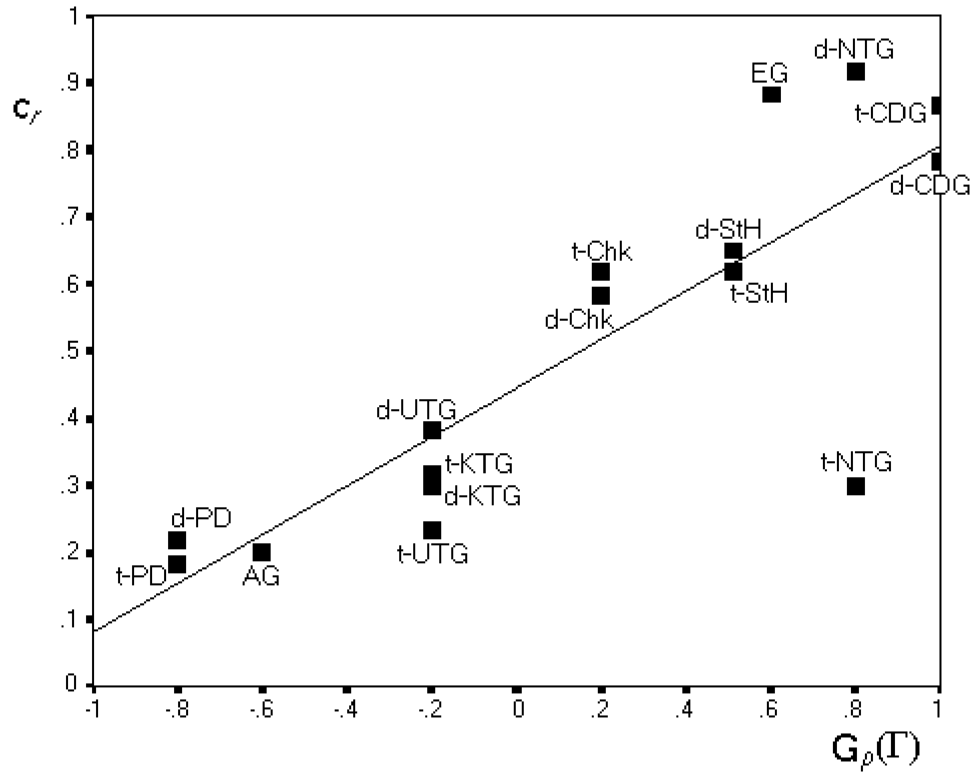




Notes: $G(\Gamma)$ and $G_\rho(\Gamma)$ are the cardinal and ordinal game harmony measures. c_Γ is the mean cooperation rate for each game. AG: Altruism Game; CDG: Coordination Game; ChK: Chicken; EG: Envy Game; PD: Prisoner's Dilemma; KTG/NTG/UTG: Kind/Needy/Unequitable Trust Game, respectively; StH: Stag-Hunt.

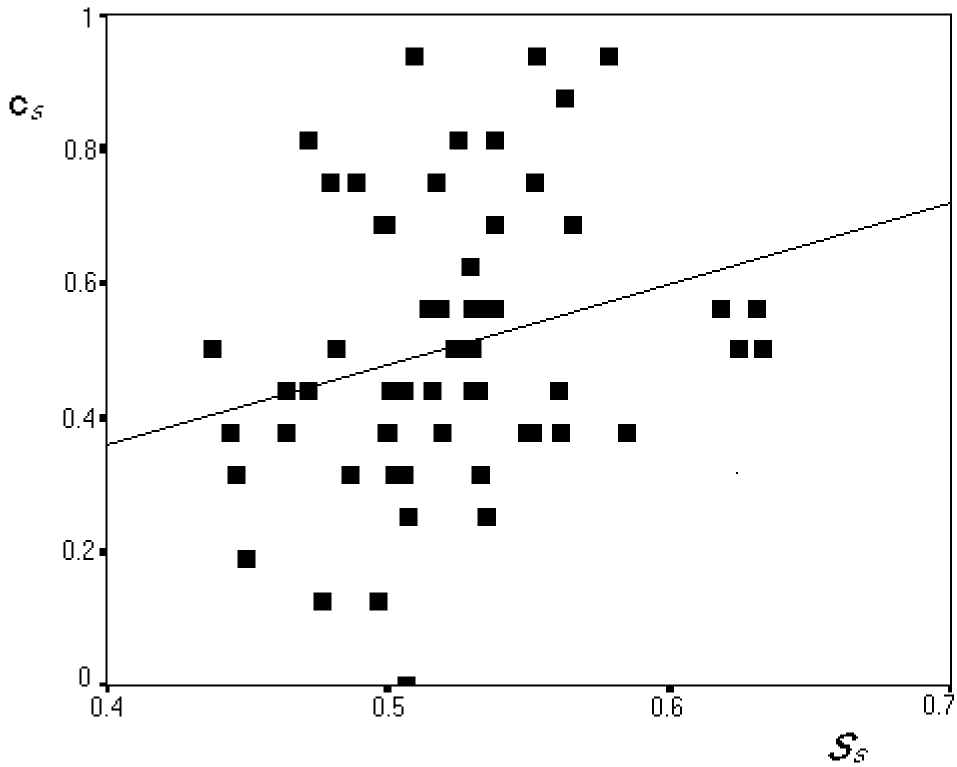
FIGURE 2. GAME HARMONY AND MEAN COOPERATION BY ROLE





Notes: see notes to Figure 1. c_r is the mean cooperation rate defined not just by game but by role. When cooperation rates can be defined both for d-players and t-players in game Γ (i.e., not for the AG and EG), d-Γ is the notation for the datapoint for d-players, and t-Γ is that for the datapoint for t-players.

FIGURE 3. MEAN SIMILARITY INDEX AND COOPERATION RATE BY SUBJECT



Notes: S_s is the mean similarity index by subject: a higher value means greater mean perceived similarity to the perfectly harmonious relative to the perfectly disharmonious game benchmark. c_s is the mean cooperation rate by subject.