ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

PERCEIVED HARMONY, SIMILARITY AND COOPERATION IN 2 x 2 GAMES: AN EXPERIMENTAL STUDY

Daniel John Zizzo and Jonathan H.W. Tan

Number 176

October 2003

Manor Road Building, Oxford OX1 3UQ

Perceived Harmony, Similarity and Cooperation in 2×2 Games: An Experimental Study

Daniel John Zizzo
BREB Unit, Department of Economics, and Christ Church College
University of Oxford

Jonathan H.W. Tan
Institute of Microeconomics
European University Viadrina

October 2003

Abstract

Game harmony is a generic game property describing how conflictual or non-conflictual the interests of players are. Simple and general game harmony measures can predict mean cooperation in 2×2 games such as the Prisoner's Dilemma, the Chicken and trust games. Two measures can be simply computed from monetary payoffs; another, the similarity index, can also be justified by theories of similarity-based reasoning. When data from Oxford and Frankfurt-Oder are disaggregated across experiments, countries and learning history, and when the similarity index is a valid measure, parsimonious regressions can explain around half of the variance in mean cooperation rates.

JEL Classification Codes: C72, C91, H41.

Keywords: game harmony, cooperation, similarity, 2×2 games, Prisoner's Dilemma.

.

Acknowledgements: We thank Michael Bacharach, David Cooper, Teng Puay Eng and participants at presentations in Oxford and Strasbourg for advice and encouragement. Thanks are also due to Jan Heufer, Jahirul Islam, Claudia Vogel, and Jessica Kaehler for experimental assistance, and to Friedel Bolle and Alexander Kritikos for checking the integrity of the English-German translations. The experimental instructions can be found in the appendix.

E-mail addresses: Zizzo: daniel.zizzo@economics.ox.ac.uk; Tan: tan@euv-frankfurt-o.de

1. Introduction

This paper presents an experimental investigation of how simple measures can be used, in different settings, to predict mean cooperation levels in well-known 2×2 games such as the Prisoner's Dilemma, the Stag-Hunt and three versions of trust games. Why bother about simple predictors for cooperation? The answer lies, in part, with the importance of predicting cooperation in a variety of settings, from social dilemmas both within and outside organizations to any transaction requiring trust, and trust fulfilment, in order for the Pareto optimal outcome to be obtained. In part, though, the answer lies in the question: because our predictors are *simple*. That is, without having to know about the distributions of cognitive abilities, beliefs, social preferences and norms in the population, given a game we may able to have a benchmark estimate of what cooperation level to expect.

Significant predictive power for our measures holds robustly in two different cultural environments (England and East Germany) and with different learning histories (three different practice stages, and different parts of the experiment performed either consecutively or about two weeks apart). The measures are based on the concept of *game harmony*, and on the expectation that higher cooperation is associated with higher harmony.

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 (Zizzo, 2003b). Consider the coordination game, constant-sum game and Prisoner's Dilemma (PD) 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, such as the PD, 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, being intuitively straightforward numerical indices, capture this "somewhere in the middle" more precisely. Zizzo's (2003a) most basic measures are an average of Pearson or Spearman correlation coefficients among payoff pairs, with higher numerical values indicating greater game harmony. In the case of 2×2 games, they reduce to simple Pearson or Spearman correlation coefficients between the payoffs of the players for each state of the world, and the Pearson correlation becomes a computationally simpler and easier to interpret version of Kelley and Thibaut's (1978) "index of correspondence".

Game harmony should be computed over the utility payoffs as perceived by economic agents, but these are not directly observable. A simple solution is to compute game harmony values purely on the basis of the monetary payoffs, and use these as predictors of cooperation. Another is to find and use a simple proxy for the degree to which the game at hand is perceived as harmonious or otherwise. We try to do both. We ask subjects not only to play the 2×2 games, but also to rate how similar each game is to a very cooperative (harmonious) and a very competitive (disharmonious) benchmark, in the guise of a pure coordination game and of a constant-sum game. To predict cooperation, we can then employ both a "basic" game harmony measure computed straight from the monetary payoff matrix, and a *similarity index* identifying how "close" a game is seen to the cooperative benchmark (where cooperation is always a good thing) relative to the competitive benchmark (where cooperation is not possible). To the extent that the similarity index has predictive power even controlling for the effect of the basic measure, it will be a proxy for how distortions in the perception of harmony in the game affect behavior. But the similarity index has also another justification. Theories of similarity-based reasoning would postulate that an agent assimilates (aspects of) the decision problem to similar one(s) that are familiar or salient, and behaves accordingly (e.g., Gilboa and Schmeidler, 2001; Markman and Moreau, 2001; Gale et al., 1995; Buschena and Zilberman, 1999). The more a game is perceived as a "like" game, the more the "like" action is likely to be chosen: a subject who assimilates a game to another game where she cooperates, such as a coordination game, is a subject who will also cooperate as a result. This provides an additional rationale for expecting a positive relationship between the mean similarity index and cooperation.

We are making no claim that game harmony is the only thing that matters for cooperation (and we shall bring a counter-example in which it does not). Yet there is much value in parsimonious predictive power (Friedman, 1953): the worth of being able to predict a significant fraction of the variance in cooperation rates may be considerable not only to economists but also to policy-makers and business managers alike willing to understand and foster cooperation and reduce conflict in a variety of settings on the basis of a minimal amount of information. Monetary payoffs are all that is needed to compute our basic game harmony measures; and a simple short questionnaire is all that is needed to elicit values of the similarity index.

Section 2 presents the experimental hypotheses and background. Sections 3 and 4 describe the methods and results of the experimental study, respectively. Section 5 contains a brief discussion and section 6 concludes.

2. Experimental Hypotheses and Background

Following Zizzo (2003b), 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 $W = W_1 \times W_2 \times \ldots \times W_n$ 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}: $W \rightarrow \Re$, 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:

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

In two-player games, this reduces to G (Γ) = $r_c(a^{\rho}_{cw}, b^{\rho}_{cw})$. 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 $x^{\rho}_{iw} = rank(x_{iw} | X_i)$, which can be mapped into rank payoff pairs $a^{\rho}_{cw}, b^{\rho}_{cw}$. Then:

$$G_\rho(\Gamma) = G_\rho(x_i, W) = \frac{1}{C}\sum_{c=1}^{C} r_c(a^{\rho}_{cw}, b^{\rho}_{cw})$$

In two-player games, this reduces to $G_\rho(\Gamma) = r_c(a^{\rho}_{cw}, b^{\rho}_{cw})$. In these games both G(Γ) and $G_\rho(\Gamma)$ 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, and should be associated with a higher likelihood of cooperation (Zizzo, 2003b).

Since utility payoffs are unobservable, one strategy is to assume that monetary payoffs reflect the utility payoffs and estimate G(Γ) and $G_\rho(\Gamma)$ directly from the matrix of monetary payoffs.

H1 (Basic Game Harmony). The mean cooperation rate c(Γ) *in relation to game* Γ *is increasing in the Pearson game harmony coefficient* $G(\Gamma)$ *and Spearman game harmony coefficient* $G_\rho(\Gamma)$.

This is not the only strategy we follow in this paper, however. Following Buschena and Zilberman (1999), in the experiment 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). Call s^Γ the similarity evaluation of a game Γ relative to the coordination game, and z^Γ the similarity evaluation relative to the constant-sum game: specifically, $s^\Gamma = z^\Gamma / (s^\Gamma + z^\Gamma)$ (if both s^Γ and z^Γ are equal to 0, s^Γ is set to 0.5). Let K be the number of separate games which are played out and which are different from the coordination and constant-sum games. Then for each subject it is possible to define the similarity index $S = \sum_{\Gamma=1}^{K} S^\Gamma / K$. Controlling for the predictive power of G(Γ) and $G_\rho(\Gamma)$, S will then be an (admittedly partial) proxy for how the subjective perception of harmony of a game differs from that entailed by its monetary payoff matrix.

Of course the interpretation of S as a proxy for perceived game harmony will hold only as long as the agent perceives indeed the constant-sum game as a disharmonious game and the coordination game as a harmonious game. An indirect indicator of the latter is cooperative play in the coordination game. Cooperative play in the coordination game is also necessary to justify a positive relationship between the similarity index and cooperation on the grounds of similarity-based reasoning.

H2 (Game Harmony and Similarity Index). If the agent plays cooperatively in the coordination game, $c(\Gamma)$ is *increasing in S, even controlling for* $G(\Gamma)$ *and* $G_\rho(\Gamma)$.

In Experiment 1, similarity evaluations were provided in a stage that followed that in which actions were chosen in the various games. One possible alternative explanation of a correlation between $c(\Gamma)$ and S is what we may label a "congruence hypothesis": perhaps subjects thought the two games similar because of cognitive dissonance (e.g., "they have to be similar, since I played a cooperative action in both") or some analogous psychological mechanism.

Our design was meant to ensure that subjects had an understanding of the games and so could give meaningful similarity evaluations. We considered the alternative of having subjects first choose similarity and then making play the games, but we believe that it would not have allowed us to discriminate H2 from the congruence hypothesis: the critic could reply that in both cases cognitive dissonance would require subjects to play similar actions in games which they both evaluated as similar to, say, the coordination game, thus spuriously increasing the correlation.

Our answer was to run a second experiment (Experiment 2) where subjects made choices in games in a first phase and then chose similarity evaluations in a second phase, taking place after about two weeks. We chose the interval to be short enough that subjects would remember what it meant playing games presented in normal form, but at the same time long enough so as to realistically reduce (or eliminate) congruence effects.

H3 (Congruence Effects). The correlation between $c(\Gamma)$ *and S should be significantly smaller in Experiment 2 than in Experiment 1.*

If congruence effects explain the correlation, a reduction of the congruence effects should reduce the correlation. As this is not predicted by H2, it provides a way to discriminate the two hypotheses.

There are two additional potential sources of variation across our sessions. First, part of our Experiment 1 was conducted in Oxford and part was conducted in Frankfurt-Oder, in (East) Germany. Second, and perhaps more seriously, we were worried that with tiny differences in the learning history and context of the decision problem the correlations with cooperation would disappear. To check this, in Experiment 1 we varied the harmony of the games used in the practice stage. While such framing differences may in fact operate by altering the connection between perceived harmony and behavior in psychologically interesting ways, we state the following hypotheses as parsimonious benchmarks.

H4 (Robustness of Basic Game Harmony). The correlations between $c(\Gamma)$ *and* $G(\Gamma)$, *and between* $c(\Gamma)$ *and* $G_\rho(\Gamma)$, *as stated by H1*, *are independent of experiment*, *venue and practice stage games*.

H5 (Robustness of Similarity Index). The correlation between $c(\Gamma)$ *and S, as stated by H2, is independent of experiment, venue and practice stage games.*

We are not aware of prior attempts to identify game harmony measures. Kelley and Thibaut (1978) devised a more complex "index of correspondence" (IC) which, unlike our measures, is applicable only to 2-player games, and which multiplies G(Γ) by an extra term, a function of the variances among the payoffs (namely, $2\sigma_1\sigma_2/(\sigma_1^2+\sigma_2^2)$). Conceptually, their measure mixes pure game harmony concerns with how much each agent stands to gain from cooperation, and so it is harder to interpret. Rapaport and Chammah (1965) devised specific cooperation indices in relation to 2-player 2 × 2 Prisoner's Dilemmas. Their measures are *ad hoc* and can only be applied to this narrow, if important, class of games. Zizzo (2003b) showed that their most successful measure appears to work as a proxy, in this specific setting, for our G(Γ), which can explain some 80% of the variance in cooperation rates in the Rapaport and Chammah's dataset. Zizzo and Tan (2003) analyzed Oxford 2 × 2 games data on mean cooperation rates at a high level of aggregation (mean cooperation rates by game and by game and player role), inclusive of sessions where a set of randomly determined games were played; Zizzo (2003a) analyzed 3 × 3 games data. Both papers estimated game harmony values relying only on the monetary payoffs in the game matrix, and found that they could be used as strong predictors of cooperation, with correlation coefficients (depending on the sample and measures used) between 0.456 and 0.962 at the high level of aggregation employed; they were also unable to find any additional predictive power from using the extra term of Kelley and Thibaut's IC formula. Effectively, these previous papers have been tests of H1 only.

3. Methods

3.1 Subjects

Experiment 1 was performed at the University of Oxford and at the University of Frankfurt-Oder. 60 subjects, mostly but not exclusively students, participated in Oxford, and 52 did so in Frankfurt-Oder. Experiment 2 was performed at the University of Frankfurt-Oder, with each subject having to participate to two phases, taking place with an interval of about two weeks; one subject who did phase 1 did not attend phase 2, but otherwise everyone did, resulting in a dataset of 23 subjects. Recruitment was mainly by the means of electronic email lists, and participation was entirely voluntary. No subject could participate more than once, or to both experiments.

3.2 Materials and Procedures

We now describe Experiment 1. Four subjects participated to each session, and were separated by partitions preventing them from seeing each other. 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.

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, both in Oxford and in Frankfurt 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. 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 sessions, respectively. Five sessions in each condition took place in Oxford; five sessions in the Medium condition, and four in the other two, took place in Frankfurt. 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 in Oxford, and 0.09 euros in Frankfurt. 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: the d-player is the truster, with a choice whether to trust (by playing top) or withhold 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 2×2 games (e.g. Rankin et al., 2000). They were used in our experiment because: 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.

Table 2 identifies unique cooperative actions on a game-to-game basis. No cooperative action is defined in the constant-sum game, nor (ignoring the payoff perturbation) is there 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*, *fulfillment*) 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 actions are obvious in the remaining cases.

Zizzo and Tan (2003) also used a rule-based procedure that treated an action as cooperative if, over the (non-payoff-perturbed) game matrix,

(a) it corresponds to both the utilitarian solution (it has the highest sum of payoffs) and the Nash Bargaining solution (it has the highest product of payoffs);

(b) it does not correspond to a strictly dominant strategy.

Criterion (b) follows the intuition that, if an action is strictly dominant, then it is best for a player to play that action irrespectively of what the other player does, and therefore it may not make much sense to talk in this case of "co-operation", interpreted as "operating with", working together with, the other player towards a common goal.

We shall refer to cooperation as determined by the game-by-game procedure and rule-based procedure as *g-cooperation* and *r-cooperation*, respectively. The r-cooperation action coincides with the g-cooperation action in all cases except three where it is undefined: it is undefined for the d-player in the EG and the NTG (the putative cooperative action corresponds to the strictly dominant strategy) and for the t-player in the Chicken game (the actions associated to the Nash Bargaining and utilitarian outcomes are different in the two cases).

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 Buschena and 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 pounds in Oxford and 18 euros in Frankfurt to get the evaluation exactly exactly right, with a penalty of 5 pounds in Oxford and 7.5 euros in Frankfurt per each point of error (subjects were paid zero for getting the guess wrong by three points or more). 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. 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 was 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.[1]

[1] 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

Payment Stage. At the end of the experiment subjects were paid the action payment, the similarity payment, plus a participation fee (4 pounds in Oxford, 5 euros in Frankfurt). Mean payments were around £ 10 in both Oxford and Frankurt, for around one hour of work.

Experiment 2. Experiment 2 was run in six sessions, and was identical to Experiment 1 except for the following. It was divided in two phases. The first phase had the practice stage (with Medium condition randomly chosen games) and the second stage, where actions were chosen in games. After about two weeks, the second phase took place. Subjects were provided some reminder instructions, in addition to questions to check their understanding. Subjects went through Stage 3 and the payment stage, which was in relation to payments earned in both phases. Financial incentives were as for the Frankfurt sessions of Experiment 1, but an additional participation fee of 5 euros in Frankfurt was paid at the end of the first phase to cover fixed costs of attendance (e.g., for travelling costs) and as a gesture of goodwill (winnings were not reported at the end of the first phase, and hence payments were deferred).

4. Results

There are three samples that we shall consider in this paper. The first is the "full sample" of all subjects. The second is the "restricted sample" of subjects (81 out of 115 in Experiment 1, 18 out of 23 in Experiment 2) who cooperated both times they played the CDG. H1 and H4 refer purely to $G(\Gamma)$ and $G_\rho(\Gamma)$, and so can be tested on both samples. Conversely, H2, H3 and H5 are cast in terms of S, and, as discussed earlier, the relationship between this and cooperation can be predicted to hold only insofar as agents do cooperate in the CDG: therefore, the restricted sample is more appropriate. The third ("intermediate") sample is inclusive of all subjects who cooperated at least once out of the two times they played the CDG (106 out of 115 in Experiment 1, everyone in Experiment 2). The intermediate sample is useful as a robustness check, and is attractive because it loses only a minimum amount of information relative to the full sample. The number of subjects removed from the sample in Experiment 1 were about the same in Oxford and Frankfurt (17 in each place to obtain the

---

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 $4.5 \times Unique + 4.5 \times [(EDP - min\,(EDP)] / [max(EDP) - min(EDP)]$.

restricted sample; 4 from the Oxford sample and 5 from the Frankfurt sample to obtain the intermediate sample).

Table 4 presents an overview of mean cooperation rates disaggregated by experiment, venue, condition, game and role, together with corresponding $G(\Gamma)$, $G_\rho(\Gamma)$, IC and S values.

(*Insert Table 4 about here)*.

The mean g-cooperation (r-cooperation) rate by subject is 0.496 (0.417) in Experiment 1, and 0.535 (0.447) in Experiment 2: the difference is statistically insignificant (e.g., in relation to g-cooperation, t (133) = 0.906, *n.s.*), a result replicated if the comparison is restricted to Medium condition sessions throughout. In Experiment 1, the g-cooperation (r-cooperation) rates were 0.453, 0.533 and 0.497 (0.367, 0.459 and 0.421) in the Low, Medium and High condition, respectively; they were 0.503 (0.487) and 0.430 (0.403) in Oxford and Frankfurt, respectively: an F test crossing Condition and Venue does not find significance for either the main effects or the interaction (e.g., in relation to g-cooperation, Condition: F (2, 106) = 1.656, *n.s.*; Venue: F (1, 106) = 0.302, *n.s.*; Condition × Venue: F (2, 106) = 0.107, *n.s.*). The stability in mean cooperation rates across venue and condition is in contrast with the large variability in cooperation rates across games: $c(\Gamma)$ moves from games such as the PD or the AG where cooperation is below 20%, to games in the middle range (around 60%) such as the ChK and the StH, to games in the high range (over 80%) such as the CDG and the EG. Figures 1 and 2 illustrate this variability in cooperation rates for each game and role, which are plotted against our game harmony proxies.

(*Insert Figures 1 and 2 about here)*.

We first consider the relationship between perceived harmony and cooperation by analysing correlation coefficients at a high level of aggregation, namely by considering the mean $c(\Gamma)$ by game and role for each experiment. Table 5 lists the correlation coefficients between $G(\Gamma)$, $G_\rho(\Gamma)$, S and IC on the one side and g-cooperation and r-cooperation rates on the other side (Figures 1 and 2 illustrate some of these correlations graphically).

(*Insert Table 5 about here)*.

$G(\Gamma)$ and $G_\rho(\Gamma)$ are strong predictors of mean cooperation rates, with correlation rates varying between 0.578 and 0.892 depending on the game harmony and cooperation measures, the correlation coefficient used, and the sample chosen (P < 0.05 or better). That is to say, basic game harmony measures alone can explain between 33 and 80% of the variance in mean cooperation rates by game and role. IC performs roughly as well (slightly better with the r-cooperation measure and slightly worse with the g-cooperation measure). As discussed earlier, the proper test with *S* is with the restricted sample. We have a medium-size effect with the g-cooperation measure, with correlation coefficients between 0.282 (*n.s.*) and 0.472 ($P < 0.05$). *S* has a good fit with the r-cooperation measure, with correlation coefficients between 0.495 ($P < 0.06$) and 0.657 ($P < 0.01$), a result robust with the other samples. Since $G(\Gamma)$ and $G_\rho(\Gamma)$ are defined for the CDG (a game with perfect harmony and high cooperation) but *S* is not (since we did not ask subjects to state how similar the CDG was to itself), part of the lesser goodness in the fit between the two sets of measures is due to the exclusion of the CDG in the set of games in relation to which *S* is a predictor.

There is no systematic tendency for the correlation between *S* and $c(\Gamma)$ to be lower in Experiment 2 relative to Experiment 1, as would be predicted by H3: to the contrary, in nine out of twelve comparisons the correlation is higher for Experiment 2 than for Experiment 1, leading to a marginal significance in the wrong direction (Wilcoxon $Z = 1.883$, $P = 0.06$, two-tailed).

We now consider mean cooperation levels disaggregated not only by game, role and experiment, but also by venue and condition (see Table 4). A regression analysis can be used to verify which patterns, and differences among patterns, are genuine and robust; it is also essential to test H2's (and H5's) prediction that *S* should have predictive power *even controlling* for basic game harmony measures. Independent variables in the general regressions are basic measure of game harmony ($G(\Gamma)$ or $G_\rho(\Gamma)$), the similarity index, and interaction terms of these variables with *Condition* (equal to 0, 1 and 2 for the Low, Medium and High condition, respectively), *Frankfurt* (equal to 1 for Frankfurt sessions, else 0) and *Experiment2* (equal to 1 for Experiment 2 sessions, else 0). These variables are regressed on either mean g-cooperation or r-cooperation $c(\Gamma)$. Since *S* is not defined over CDG, the regression analysis omits CDG observations (this is likely to reduce overall explanatory power, since $G(\Gamma)$ and $G_\rho(\Gamma)$ are good predictors for $c(\Gamma)$). H1 predicts a positive coefficient on the $G(\Gamma)$ or $G_\rho(\Gamma)$ term. H2 predicts a positive coefficient on *S* even when the $G(\Gamma)$ or $G_\rho(\Gamma)$ term is also included. H3 predicts that the coefficient on $S \times$ *Experiment2* term should be significant and negative. H4 predicts that the coefficients on the interaction terms with

$G(\Gamma)$ or $G_\rho(\Gamma)$ should be insignificant, and H5 predicts the same in relation to the interaction terms with *S*. Details of the regression analysis are in Table 6.

(*Insert Table 6 about here*).

The R^2 is lowest but still considerable for the full sample (between 0.348 and 0.394), and highest for the restricted sample (between 0.423 and 0.632). It is always higher for g-cooperation regressions due to the better fit of $G(\Gamma)$ and $G_\rho(\Gamma)$ in this dataset. Nevertheless, there is unequivocal support for H1 in both g-cooperation and r-cooperation regressions, with the basic game harmony term always statistically significant at $P < 0.05$ or better. Furthermore, the interaction terms involving $G(\Gamma)$ or $G_\rho(\Gamma)$ are always insignificant, providing support for H4. Basic game harmony measures are strong predictors of cooperation, independently of the specific measure of harmony or of cooperation used, of whether the experiment is run in England or Germany, in one or two phases, and of the harmony of practice stage games.

The correct test for H2 is with the restricted sample: the coefficient on *S* is always significant, at the 0.1 level when coupled with g-cooperation and $G(\Gamma)$, at the 0.05 level when coupled with g-cooperation and $G_\rho(\Gamma)$, and at the 0.01 level when coupled with r-cooperation (and either $G(\Gamma)$ or $G_\rho(\Gamma)$). The worse fit with the full sample is to be expected due to the sure inappropriateness of *S* with subjects who never cooperate in the CDG. Even so, in the full sample *S* remains statistically significant at the 0.01 level with r-cooperation measures, while in the intermediate sample the fit is almost as good as in the restricted sample. Overall the evidence is favourable to H2, especially (but not exclusively) where r-cooperation is to be predicted: whenever its use is appropriate, the similarity index is a significant predictor even controlling for $G(\Gamma)$ and $G_\rho(\Gamma)$.

Similarity × *Experiment2* is always statistically insignificant, providing support for H5 and no support for H3. *Similarity* × *Frankfurt* is also insignificant, as for H5. The evidence is more equivocal for *Similarity* × *Condition*, which is significant at the 0.1 or 0.05 level in the restricted sample (only). It implies that (whenever the CDG was seen as a cooperative benchmark) higher practice stage harmony may have made assimilation to the CDG benchmark more powerful in raising cooperation levels. Overall, there is no support for H3 and partial support for H5. The correlation between similarity index and cooperation does not appear to be spuriously raised by congruence effects, and the effectiveness of the similarity index appears independent of venue and of whether the experiment is run in one or two phases. While the context may affect the effectiveness of the similarity index, the coefficients on *Similarity* × *Condition* are considerably smaller than those on *Similarity*, suggesting a degree of robustness.

5. Discussion

We analyzed the predictive power of simple game harmony measures in relation to a sample of mostly well-known and important 2 × 2 games. Two basic measures can be computed relying purely on the monetary payoffs of the games. Another, the similarity index, can be obtained by a small set of verbal responses, and verifies how "close" a game is perceived to a coordination game (a perfectly harmonious benchmark) relative to a constant-sum game (a perfectly disharmonious benchmark).

Our basic measures have significant predictive power when different measures of cooperation and of correlation are used, with different samples, in two different countries (England and East Germany), with different learning histories and whether the experiment is in one or two temporally separated phases. The similarity index is also an effective (if individually less powerful) measure, especially when subjects for whom it is not a valid proxy for game harmony are excluded, and especially when one of the two cooperation measures is used. It is a robust measure across countries and with different learning histories, although having previously faced games with higher game harmony may slightly augment its predictive power. Its association with cooperation holds even controlling for basic game harmony measures, and can be interpreted as evidence for similarity-based reasoning and/or as a proxy for how the way games are perceived by the agent differs from what is just entailed by monetary payoffs. Conversely, it cannot be interpreted as spuriously induced by some psychological mechanism producing a congruence of similarity evaluations with the actions undertaken, as we ran an experiment specifically to test this hypothesis, and failed to find support for it.

When the data is disaggregated across experiments (one or two phases), countries and learning history (practice stage condition), and when the similarity index is a valid measure, regressions of mean cooperation rate on basic harmony measures and on the similarity index, plus an interaction term with practice stage harmony, can explain between 42 and 63% of the variance. With the full sample (inclusive of observations where the similarity index is not a valid measure), the explanatory power with just the basic measure and the similarity index is for at least 35% of the variance. When the data is at a higher level of aggregation (that is, aggregated by game and game role), basic game harmony measures alone can explain between 36 and 73% of the variance in mean cooperation rates by game and role.

We are making no claim, of course, that game harmony is the only thing that matters for cooperation. For example, 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 7 contains a pair of games exemplifying this limitation.

(*Insert Table 7 about here)*.

What we claim is that game harmony measures can be used as predictors of cooperation, at least in our dataset of 2×2 games. Zizzo and Tan (2003) and Zizzo (2003a) respectively show the power of basic measures of game harmony in the context of randomly determined 2×2 games and of 3×3 games; further work could be made to test the similarity index beyond our dataset. Predicting cooperation is important for many real-world situations, for example to understand behavior in organizational structures, to design environmental policies, and to define terms in bilateral trade negotiations. Our measures are *simple*, as they rely on only minimal informational requirements, and *general*, as their applicability goes well beyond 2-player games.

6. Conclusion

Simple and general game harmony measures can be used to predict cooperation in a sample of 2×2 games such as the Prisoner's Dilemma, the Chicken and three version of trust games. Two basic measures can be computed relying purely on the monetary payoffs. When controlling for them, another measure, the similarity index, proxies for the degree to which the perception of harmony of the game differs from that implied by the monetary payoffs; its association with cooperation can also be justified by theories of similarity-based reasoning. When our data from experiments in Oxford and Frankfurt-Oder is disaggregated across experiments, countries and learning history, and when the similarity index is a valid measure, regressions of mean cooperation rate on basic harmony measures and on the similarity index, plus an interaction term showing some history-dependence, can explain between 42 and 63% of the variance in mean cooperation rates. When the data is at a higher level of aggregation, basic game harmony measures alone can explain between 36 and 73% of the variance in mean cooperation rates. With minimal knowledge required about the distributions of cognitive abilities, beliefs or social preferences and norms in the population, we may be able to have a robust benchmark estimate of what cooperation level to expect in a given game, before factoring in additional strategic considerations.

Appendix

Experimental Instructions

Notations: *[.] *denotes text relevant exclusively to Experiment 1, and* **[.] *denotes text relevant exclusively to Experiment 2. When monetary amounts are mentioned, they are done so in terms of "x pounds/ y euros", although of course only pounds were mentioned in the Oxford sessions, and only euros in the Frankfurt sessions. These notations are not found in the original instructions.*

The experimental instructions used in Germany were German translations of the English version.

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/5 euros for participation.]

**[You are about to participate in an experiment on decision-making. The experiment will be conducted in two parts, the first today (Part 1) and the next in 10-14 days time (Part 2).

Today you will do Stage 1 and 2, and receive 5 euros for participation at the end. In Part 2, you will do Stage 3 and 4. There are four stages overall.

Stage 1 (the Practice Stage) is for practice only, while in Stage 4 (the Payment Stage) you are paid whatever amount you have earned in Stages 2 and 3, plus additional 5 euros for participation. You must participate to both parts of the experiment to be paid anything more than the 5 euros at the end of today's Part 1.]

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 four 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 (*amount*) points and your coparticipant gets (*amount*); if you choose B and your coparticipant chooses B, you get (*amount*) points and your coparticipant gets (*amount*); finally, if you choose B and your coparticipant chooses A, you get (*amount*) points and your coparticipant gets (*amount*). 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.

Instructions for Stage 2

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/0.09 euros in the Payment Stage* (so, for example, 100 points are worth 6 pounds/9 euros).

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

**[Welcome to Part 2 of this experiment. In Part 1 you did Stage 1 and 2, and now you are about to do Stage 3.

Reminder About Decision Tables

You will recall that in Part 1 of the experiment you chose actions in relation to Decision Tables. Decision Tables consist of eight numbers arranged in two rows and two columns.

Consider the Decision Table in the top part of the screen. Assume that you and your coparticipant had to choose actions in relation to this Decision Table. You and your coparticipant would 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 four 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 chose A and your coparticipant chose A, you would get (*amount*) points and your coparticipant would get (*amount*); if you chose B and your coparticipant chose B, you would get (*amount*) points and your coparticipant would get (*amount*); finally, if you chose B and your coparticipant chose A, you would get (*amount*) points and your coparticipant would get (*amount*). The point numbers in all cells are always between 0 and 100.

Similarity Evaluations]

In Stage 3 you are asked to ***evaluate how similar two Decision Tables are*** (a reminder about what Decision Tables is provided below). 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/18 euros**. 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/7.5 euros**. 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 /euros)**
0	**12/18**
1	**7/10.5**
2	**2/3**
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.

References

Buschena, David and Zilberman, David (1999). “Testing the Effects of Similarity on Risky Choice: Implications for Violations of Expected Utility,” *Theory and Decision*, 46, pp. 251-276.

Friedman, Milton (1953). *Essays in Positive Economics*. Chicago: Chicago University Press.

Gale, John, Binmore, Kenneth G. and Samuelson, Larry (1995), “Learning to Be Imperfect: The Ultimatum Game,” *Games and Economic Behavior*, 8, pp. 856-890.

Gilboa, Itzhak, and Schmeidler, David (2001). *A Theory of Case-Based Decisions*. Cambridge: Cambridge University Press.

Kelley, Harold H. and Thibaut, John W. (1978). *Interpersonal Relations: A Theory of Interdependence*. New York: Wiley.

Markman, Arthur B. and Moreau, C. Page (2001). “Analogy and Analogical Comparison in Choice.” In *The Analogical Mind: Perspectives from Cognitive Science*, edited by Derdre Gentner, Keith J. Holyoak and Boicho N. Kokinov. Cambridge (Mass.) and London: MIT Press, pp. 363-399.

Rankin, Frederick W., Van Huyck, John B. and Battalio, Raymond C. (2000). “Strategic Similarity and Emergent Conventions: Evidence from Similar Stag Hunt Games,” *Games and Economic Behavior*, 32, pp. 315-337.

Rapoport, Anatol and Chammah, Albert M. (1965). *Prisoner's Dilemma: A Study in Conflict and Cooperation*. Ann Arbor: University of Michigan Press.

Zizzo, D.J. (2003a). “Game Perception and Harmony in 3 × 3 Games,” Oxford University Department of Economics Discussion Paper 152.

Zizzo, D.J. (2003b). “Harmony of Games in Normal Form,” Oxford University Department of Economics Discussion Paper 150.

Zizzo, D.J. and Tan, J.H.W. (2003). “Game Harmony as a Predictor of Cooperation in 2×2 Games,” Oxford University Department of Economics Discussion Paper 151.

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 EXPERIMENT A 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**

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. MEAN COOPERATION RATES

Experiment 1												
Oxford, Low Condition												
	Full Sample		Restr. Sample		Interm. Sample					Full	Restr.	Interm.
Game	g-c (Γ)	r-c (Γ)	g-c (Γ)	r-c (Γ)	g-c (Γ)	r-c (Γ)	G(Γ)	Gr(Γ)	IC	S	S	S
d-PD	0.15	0.15	0.188	0.188	0.176	0.176	-0.817	-0.8	-0.817	0.571	0.566	0.562
d-CDG	0.8	0.8	1	1	0.941	0.941	1	1	0.999			
d-EG	0.9		0.938		0.941		0.98	0.6	0.468	0.454	0.436	0.44
d-AG	0.1	0.1	0.125	0.125	0.118	0.118	-0.98	-0.6	-0.468	0.475	0.444	0.476
d-KTG	0.2	0.2	0.188	0.188	0.176	0.176	0.066	-0.2	0.065	0.608	0.663	0.624
d-UTG	0.4	0.4	0.438	0.438	0.412	0.412	0.066	-0.2	0.065	0.459	0.483	0.455
d-NTG	0.9		0.938		0.941		0.503	0.8	0.5	0.498	0.519	0.494
d-St-H	0.65	0.65	0.813	0.813	0.765	0.765	0.488	0.513	0.488	0.623	0.637	0.647
d-Chk	0.6	0.6	0.625	0.625	0.647	0.647	0.16	0.2	0.16	0.56	0.577	0.57
t-PD	0.15	0.15	0.188	0.188	0.176	0.176	-0.817	-0.8	-0.817	0.504	0.5	0.5
t-CDG		0.1	1	1	1	1	1	1	0.999	0.522		
t-EG				0.125		0.118	0.98	0.6	0.468	0.481	0.535	0.53
t-KTG	0.2	0.2	0.25	0.25	0.235	0.235	0.066	-0.2	0.065	0.549	0.534	0.562
t-UTG	0.15	0.15	0.188	0.188	0.176	0.176	0.066	-0.2	0.065	0.522	0.532	0.53
t-NTG	0.15	0.15	0.188	0.188	0.176	0.176	0.503	0.8	0.5	0.515	0.563	0.53
t-St-H	0.65	0.65	0.688	0.688	0.647	0.647	0.488	0.513	0.488	0.565	0.611	0.575
t-Chk	0.65		0.625		0.647		0.16	0.2	0.16	0.547	0.54	0.559
Oxford, Medium Condition												
	Full Sample		Restr. Sample		Interm. Sample					Full	Restr.	Interm.
Game	g-c (Γ)	r-c (Γ)	g-c (Γ)	r-c (Γ)	g-c (Γ)	r-c (Γ)	G(Γ)	Gr(Γ)	IC	S	S	S
d-PD	0.2	0.2	0.2	0.2	0.211	0.211	-0.817	-0.8	-0.817	0.536	0.558	0.545
d-CDG	0.8	0.8	1	1	0.842	0.842	1	1	0.999			
d-EG	0.95		0.933		0.947		0.98	0.6	0.468	0.518	0.522	0.519
d-AG	0.25	0.25	0.133	0.133	0.211	0.211	-0.98	-0.6	-0.468	0.468	0.466	0.48
d-KTG	0.25	0.25	0.267	0.267	0.263	0.263	0.066	-0.2	0.065	0.603	0.635	0.609
d-UTG	0.45	0.45	0.467	0.467	0.474	0.474	0.066	-0.2	0.065	0.474	0.467	0.478
d-NTG	0.9		0.867		0.895		0.503	0.8	0.5	0.469	0.483	0.466
d-St-H	0.65	0.65	0.667	0.667	0.684	0.684	0.488	0.513	0.488	0.567	0.54	0.572
d-Chk	0.55	0.55	0.533	0.533	0.579	0.579	0.16	0.2	0.16	0.575	0.571	0.572
t-PD	0.15	0.15	0.133	0.133	0.158	0.158	-0.817	-0.8	-0.817	0.516	0.497	0.516
t-CDG		0.55	1	1	0.947	0.947	1	1	0.999	0.474		
t-EG				0.6		0.579	0.98	0.6	0.468	0.438	0.47	0.48
t-KTG	0.4	0.4	0.4	0.4	0.421	0.421	0.066	-0.2	0.065	0.485	0.488	0.497
t-UTG	0.2	0.2	0.2	0.2	0.211	0.211	0.066	-0.2	0.065	0.443	0.419	0.446
t-NTG	0.4	0.4	0.333	0.333	0.368	0.368	0.503	0.8	0.5	0.483	0.47	0.484
t-St-H	0.75	0.75	0.8	0.8	0.737	0.737	0.488	0.513	0.488	0.588	0.611	0.6
t-Chk	0.7		0.667		0.684		0.16	0.2	0.16	0.56	0.566	0.556
Oxford, High Condition												
	Full Sample		Restr. Sample		Interm. Sample					Full	Restr.	Interm.
Game	g-c (Γ)	r-c (Γ)	g-c (Γ)	r-c (Γ)	g-c (Γ)	r-c (Γ)	G(Γ)	Gr(Γ)	IC	S	S	S
d-PD	0.3	0.3	0.333	0.333	0.3	0.3	-0.817	-0.8	-0.817	0.502	0.505	0.502
d-CDG	0.8		1	1	0.75	0.75	1	1	0.999	0.513		
d-EG	0.25	0.25	0.917		0.8		0.98	0.6	0.468	0.49	0.569	0.513
d-AG	0.45	0.45	0.25	0.25	0.25	0.25	-0.98	-0.6	-0.468	0.61	0.493	0.49
d-KTG	0.3	0.3	0.667	0.667	0.45	0.45	0.066	-0.2	0.065	0.505	0.608	0.61
d-UTG	0.95		0.333	0.333	0.3	0.3	0.066	-0.2	0.065	0.46	0.515	0.505
d-NTG	0.65	0.65	1		0.95		0.503	0.8	0.5	0.568	0.45	0.46
d-St-H	0.6	0.6	1	1	0.65	0.65	0.488	0.513	0.488	0.588	0.61	0.568
d-Chk	0.25	0.25	0.583	0.583	0.6	0.6	0.16	0.2	0.16	0.529	0.551	0.588
t-PD	0.85	0.85	0.333	0.333	0.25	0.25	-0.817	-0.8	-0.817		0.514	0.529
t-CDG		0.5	1	1	0.85	0.85	1	1	0.999	0.437		
t-EG				0.667		0.5	0.98	0.6	0.468	0.453	0.473	0.437
t-KTG	0.35	0.35	0.333	0.333	0.35	0.35	0.066	-0.2	0.065	0.61	0.601	0.61
t-UTG	0.35	0.35	0.417	0.417	0.35	0.35	0.066	-0.2	0.065	0.452	0.498	0.452
t-NTG	0.35	0.35	0.583	0.583	0.35	0.35	0.503	0.8	0.5	0.527	0.503	0.527
t-St-H	0.45	0.45	0.583	0.583	0.45	0.45	0.488	0.513	0.488	0.561	0.566	0.561
t-Chk	0.5		0.5		0.5		0.16	0.2	0.16	0.54	0.574	0.54

Frankfurt-Oder, Low Condition

Game	Full Sample g-c(Γ)	Full Sample r-c(Γ)	Restr. Sample g-c(Γ)	Restr. Sample r-c(Γ)	Interm. Sample g-c(Γ)	Interm. Sample r-c(Γ)	G(Γ)	Gr(Γ)	IC	Full S	Restr. S	Interm. S
d-PD	0.125	0.125	0.2	0.2	0.143	0.143	-0.817	-0.8	-0.817	0.557	0.536	0.542
d-CDG	0.625	0.625	1	1	0.929	0.929	1	1	0.999	0.563		
d-EG	0.813	0.813	1		1		0.98	0.6	0.468		0.519	0.504
d-AG	0.938		0.1	0.1	0.071	0.071	-0.98	-0.6	-0.468	0.482	0.512	0.53
d-KTG	0.063	0.063	0.3	0.3	0.286	0.286	0.066	-0.2	0.065	0.522	0.637	0.622
d-UTG	0.313	0.313	0.4	0.4	0.5	0.5	0.066	-0.2	0.065	0.615	0.441	0.488
d-NTG	0.5	0.5	0.9		0.929		0.503	0.8	0.5	0.465	0.487	0.457
d-St-H	0.875		0.6	0.6	0.5	0.5	0.488	0.513	0.488	0.469	0.523	0.517
d-Chk	0.438	0.438	0.7	0.7	0.714	0.714	0.16	0.2	0.16	0.528	0.59	0.572
t-PD	0	0	0	0	0	0	-0.817	-0.8	-0.817	0.475	0.499	0.46
t-CDG	0.313		1	1	0.786	0.786	1	1	0.999	0.591		
t-EG		0.313		0.4		0.286	0.98	0.6	0.468	0.503	0.513	0.508
t-KTG			0.4	0.4	0.286	0.286	0.066	-0.2	0.065	0.497	0.608	0.583
t-UTG	0.25	0.25	0.2	0.2	0.143	0.143	0.066	-0.2	0.065	0.598	0.436	0.427
t-NTG	0.125	0.125	0.2	0.2	0.143	0.143	0.503	0.8	0.5	0.428	0.518	0.461
t-St-H	0.188	0.188	0.9	0.9	0.714	0.714	0.488	0.513	0.488	0.484	0.517	0.509
t-Chk	0.688	0.688	0.2		0.357		0.16	0.2	0.16	0.505	0.591	0.577

Frankfurt-Oder, Medium Condition

Game	Full Sample g-c(Γ)	Full Sample r-c(Γ)	Restr. Sample g-c(Γ)	Restr. Sample r-c(Γ)	Interm. Sample g-c(Γ)	Interm. Sample r-c(Γ)	G(Γ)	Gr(Γ)	IC	Full S	Restr. S	Interm. S
d-PD	0.043	0.043	0.056	0.056	0.043	0.043	-0.817	-0.8	-0.817	0.517	0.54	0.517
d-CDG	0.783	0.783	1	1	0.783	0.783	1	1	0.999			
d-EG	0.913		1		0.913		0.98	0.6	0.468	0.512	0.53	0.512
d-AG	0.174	0.174	0.167	0.167	0.174	0.174	-0.98	-0.6	-0.468	0.476	0.461	0.476
d-KTG	0.304	0.304	0.333	0.333	0.304	0.304	0.066	-0.2	0.065	0.614	0.638	0.614
d-UTG	0.435	0.435	0.444	0.444	0.435	0.435	0.066	-0.2	0.065	0.528	0.517	0.528
d-NTG	1		1		1		0.503	0.8	0.5	0.537	0.512	0.537
d-St-H	0.696	0.696	0.778	0.778	0.696	0.696	0.488	0.513	0.488	0.567	0.559	0.567
d-Chk	0.696	0.696	0.778	0.778	0.696	0.696	0.16	0.2	0.16	0.539	0.554	0.539
t-PD	0.174	0.174	0.167	0.167	0.174	0.174	-0.817	-0.8	-0.817	0.519	0.528	0.519
t-CDG		0.261	1	1	1	1	1	1	0.999	0.553		
t-EG				0.278		0.261	0.98	0.6	0.468	0.511	0.567	0.553
t-KTG	0.478	0.478	0.556	0.556	0.478	0.478	0.066	-0.2	0.065	0.59	0.562	0.59
t-UTG	0.261	0.261	0.222	0.222	0.261	0.261	0.066	-0.2	0.065	0.405	0.382	0.405
t-NTG	0.348	0.348	0.389	0.389	0.348	0.348	0.503	0.8	0.5	0.474	0.471	0.474
t-St-H	0.609	0.609	0.722	0.722	0.609	0.609	0.488	0.513	0.488	0.605	0.618	0.605
t-Chk	0.652		0.667		0.652		0.16	0.2	0.16	0.583	0.599	0.583

Frankfurt-Oder, High Condition

Game	Full Sample g-c(Γ)	Full Sample r-c(Γ)	Restr. Sample g-c(Γ)	Restr. Sample r-c(Γ)	Interm. Sample g-c(Γ)	Interm. Sample r-c(Γ)	G(Γ)	Gr(Γ)	IC	Full S	Restr. S	Interm. S
d-PD	0	0	0	0	0	0	-0.817	-0.8	-0.817	0.536	0.555	0.532
d-CDG	0.625	0.625	1	1	0.714	0.714	1	1	0.999			
d-EG	1		1		1		0.98	0.6	0.468	0.583	0.59	0.595
d-AG	0.125	0.125	0.111	0.111	0.143	0.143	-0.98	-0.6	-0.468	0.503	0.554	0.508
d-KTG	0.313	0.313	0.444	0.444	0.357	0.357	0.066	-0.2	0.065	0.555	0.531	0.533
d-UTG	0.438	0.438	0.667	0.667	0.429	0.429	0.066	-0.2	0.065	0.49	0.51	0.489
d-NTG	0.938		1		1		0.503	0.8	0.5	0.555	0.592	0.561
d-St-H	0.563	0.563	0.889	0.889	0.643	0.643	0.488	0.513	0.488	0.514	0.54	0.501
d-Chk	0.563	0.563	0.556	0.556	0.571	0.571	0.16	0.2	0.16	0.524	0.531	0.519
t-PD	0.125	0.125	0.111	0.111	0.143	0.143	-0.817	-0.8	-0.817	0.496	0.486	0.486
t-CDG		0.25	1	1	0.929	0.929	1	1	0.999	0.525		
t-EG				0.222		0.214	0.98	0.6	0.468	0.548	0.501	0.508
t-KTG	0.25	0.25	0.333	0.333	0.214	0.214	0.066	-0.2	0.065	0.546	0.506	0.529
t-UTG	0.25	0.25	0.333	0.333	0.214	0.214	0.066	-0.2	0.065	0.456	0.42	0.449
t-NTG	0.188	0.188	0.222	0.222	0.214	0.214	0.503	0.8	0.5	0.547	0.556	0.556
t-St-H	0.75	0.75	1	1	0.857	0.857	0.488	0.513	0.488	0.577	0.587	0.566
t-Chk	0.75		0.667		0.786		0.16	0.2	0.16	0.585	0.595	0.582

Experiment 2 (Frankfurt-Oder, Medium)												
	Full Sample		Restr. Sample		Interm. Sample					Full	Restr.	Interm.
Game	g-*c* (Γ)	r-*c* (Γ)	g-*c* (Γ)	r-*c* (Γ)	g-*c* (Γ)	r-*c* (Γ)	G(Γ)	Gr(Γ)	IC	*S*	*S*	*S*
d-PD	0.15	0.15	0.125	0.125	0.158	0.158	-0.817	-0.8	-0.817	0.549	0.544	0.541
d-CDG	0.8	0.8	1	1	0.842	0.842	1	1	0.999			
d-EG	0.95		0.938		0.947		0.98	0.6	0.468	0.526	0.512	0.501
d-AG	0.25	0.25	0.188	0.188	0.263	0.263	-0.98	-0.6	-0.468	0.526	0.497	0.501
d-KTG	0.45	0.45	0.5	0.5	0.474	0.474	0.066	-0.2	0.065	0.611	0.596	0.59
d-UTG	0.45	0.45	0.438	0.438	0.474	0.474	0.066	-0.2	0.065	0.508	0.477	0.482
d-NTG	0.9		0.875		0.895		0.503	0.8	0.5	0.51	0.511	0.51
d-St-H	0.7	0.7	0.813	0.813	0.737	0.737	0.488	0.513	0.488	0.577	0.644	0.608
d-Chk	0.6	0.6	0.5	0.5	0.579	0.579	0.16	0.2	0.16	0.489	0.517	0.514
t-PD	0.2	0.2	0.25	0.25	0.211	0.211	-0.817	-0.8	-0.817	0.514	0.52	0.516
t-CDG		0.15	1	1	1	1	1	1	0.999	0.516		
t-EG				0.125		0.158	0.98	0.6	0.468	0.442	0.548	0.543
t-KTG	0.45	0.45	0.438	0.438	0.474	0.474	0.066	-0.2	0.065	0.611	0.623	0.625
t-UTG	0.15	0.15	0.188	0.188	0.158	0.158	0.066	-0.2	0.065	0.327	0.349	0.344
t-NTG	0.3	0.3	0.313	0.313	0.316	0.316	0.503	0.8	0.5	0.46	0.453	0.468
t-St-H	0.75	0.75	0.875	0.875	0.789	0.789	0.488	0.513	0.488	0.659	0.658	0.641
t-Chk	0.5		0.5		0.526		0.16	0.2	0.16	0.637	0.618	0.617

Restr. and interm. sample refer to the restricted and the intermediate sample, respectively. Full *S*, restr. *S* and interm. *S* refer to the similarity index in the full, restricted and intermediate sample, respectively. g-*c*(Γ) and r-*c*(Γ) refer to g-cooperation and r-cooperation, respectively.

TABLE 5. CORRELATION BETWEEN PERCEIVED HARMONY AND COOPERATION

	Full Sample		Restricted Sample		Intermediate Sample	
	Pearson correlation coefficients, Experiment 1					
	g-*c* (Γ)	r-*c* (Γ)	g-*c* (Γ)	r-*c* (Γ)	g-*c* (Γ)	r-*c* (Γ)
G(Γ)	0.825***	0.771***	0.872***	0.767***	0.836***	0.746***
$G_\rho(\Gamma)$	0.789***	0.813***	0.867***	0.767***	0.839***	0.787***
IC	0.789***	0.813***	0.849***	0.841***	0.805***	0.823***
S	0.171	0.515**	0.282	0.549**	0.186	0.533**
	Pearson correlation coefficients, Experiment 2					
	g-*c* (Γ)	r-*c* (Γ)	g-*c* (Γ)	r-*c* (Γ)	g-*c* (Γ)	r-*c* (Γ)
G(Γ)	0.844***	0.872***	0.854***	0.624***	0.834***	0.606***
$G_\rho(\Gamma)$	0.849***	0.765***	0.834***	0.683***	0.821***	0.671***
IC	0.824***	0.816***	0.837***	0.740***	0.815***	0.732***
S	0.327	0.447*	0.472**	0.657***	0.403*	0.624**
	Spearman correlation coefficients, Experiment 1					
	g-*c* (Γ)	r-*c* (Γ)	g-*c* (Γ)	r-*c* (Γ)	g-*c* (Γ)	r-*c* (Γ)
G(Γ)	0.855***	0.745***	0.892***	0.789***	0.853***	0.774***
$G_\rho(\Gamma)$	0.828***	0.754***	0.853***	0.767***	0.839***	0.778***
IC	0.802***	0.803***	0.838***	0.807***	0.821***	0.818***
S	0.248	0.469*	0.374*	0.613**	0.288	0.469*
	Spearman correlation coefficients, Experiment 2					
	g-*c* (Γ)	r-*c* (Γ)	g-*c* (Γ)	r-*c* (Γ)	g-*c* (Γ)	r-*c* (Γ)
G(Γ)	0.870***	0746***	0.864***	0.580**	0.845***	0.578**
$G_\rho(\Gamma)$	0.855***	0.764***	0.837***	0.602***	0.822***	0.614**
IC	0.832***	0.82***	0.819***	0.692***	0.802***	0.700***
S	0.355	0.531**	0.399*	0.496*	0.285	0.489*

TABLE 6. REGRESSIONS ON G-COOPERATION (G-$C(\Gamma)$) AND R-COOPERATION (R-$C(\Gamma)$) RATE

	Reg. with $G(\Gamma)$ - Full Sample				Reg. with $G_\rho(\Gamma)$ - Full Sample			
	g-*c* (Γ)		r-*c* (Γ)		g-*c* (Γ)		r-*c* (Γ)	
	t	S.E.	t	S.E.	t	S.E.	t	S.E.
G(Γ)	0.24***	0.084	0.15**	0.068	0.287**	0.093	0.169**	0.081
S	0.46	0.404	1.117***	0.258	0.5*	0.384	1.19***	0.252
G(Γ) × *Frankfurt*	-0.003	0.1	0.054	0.064	0.002	0.098	0.037	0.072
G(Γ) × *Experiment*	0.086	0.104	-0.06	0.11	0.069	0.122	-0.068	0.117
G(Γ) × *Condition*	0.032	0.07	-0.014	0.04	0.017	0.064	-0.018	0.046
S × *Frankfurt*	-0.014	0.098	-0.063	0.074	-0.02	0.09	-0.053	0.073
S × *Experiment*	0.088	0.099	0.12	0.101	0.098	0.092	0.112	0.103
S × *Condition*	0.041	0.066	0.068	0.045	0.098	0.061	0.063	0.045
Constant	0.159	0.219	-0.275**	0.131	0.14	0.214	-0.31**	0.131
R^2	0.373		0.369		0.394		0.384	
	F test for linear restrictions (all *n.s.*):							
	F(6, 89) = 0.41		F(6, 75) = 1.32		F(6, 89) = 0.34		F(6, 75) = 1.30	
	t	S.E.	t	S.E.	t	S.E.	t	S.E.
G(Γ)	0.045*	6.27	0.156***	0.03	0.294***	0.044	0.163***	0.032
S	0.545*	0.389	1.2***	0.259	0.604*	0.375	1.243***	0.255
Constant	0.139	0.214	-0.292**	0.134	0.11	0.208	-0.312**	0.135
R^2	0.358		0.352		0.381		0.355	
	Reg. with G(Γ) - Restricted Sample				Reg. with $G_\rho(\Gamma)$ - Restricted Sample			
	g-*c* (Γ)		r-*c* (Γ)		g-*c* (Γ)		r-*c* (Γ)	
	t	S.E.	t	S.E.	t	S.E.	t	S.E.
G(Γ)	0.361***	0.061	0.158**	0.078	0.379***	0.097	0.201**	0.097
S	0.474*	0.361	1.161***	0.321	0.567**	0.312	1.115***	0.297
G(Γ) × *Frankfurt*	0.09	0.061	0.051	0.081	0.062	0.096	0.042	0.096
G(Γ) × *Experiment*	-0.075	0.076	-0.115	0.137	-0.071	0.121	-0.099	0.134
G(Γ) × *Condition*	0	0.04	0.04	0.053	0	0.063	0.031	0.066
S × *Frankfurt*	-0.008	0.075	-0.018	0.088	-0.005	0.074	-0.015	0.091
S × *Experiment*	0.025	0.083	0.002	0.12	0.02	0.085	-0.007	0.124
S × *Condition*	0.093*	0.049	0.115**	0.057	0.093*	0.049	0.117*	0.06
Constant	0.165	0.187	-0.276*	0.16	0.123	0.165	-0.241	0.154
R^2	0.632		0.437		0.609		0.455	
	F test for linear restrictions (all *n.s.*):							
	F(5, 89) = 0.62		F(5, 75) = 0.34		F(5, 89) = 0.28		F(5, 75) = 0.23	
	t	S.E.	t	S.E.	t	S.E.	t	S.E.
G(Γ)	0.402***	0.03	0.211***	0.039	0.045	9.01	0.045*	5.42
S	0.469*	0.336	1.11***	0.297	0.574**	0.289	1.088***	0.275
S × *Condition*	0.094*	0.051	0.117**	0.058	0.094*	0.052	0.115**	0.057
Constant	0.167	0.178	-0.256*	0.153	0.119	0.159	-0.231	0.147
R^2	0.625		0.423		0.606		0.447	
	Reg. with G(Γ) - Intermediate Sample				Reg. with $G_\rho(\Gamma)$ - Intermediate Sample			
	g-*c* (Γ)		r-*c* (Γ)		g-*c* (Γ)		r-*c* (Γ)	
	t	S.E.	t	S.E.	t	S.E.	t	S.E.
G(Γ)	0.365***	0.06	0.157**	0.072	0.378***	0.095	0.19**	0.085
S	0.47	0.383	1.17***	0.286	0.585**	0.323	1.156***	0.27
G(Γ) × *Frankfurt*	0.082	0.061	0.035	0.067	0.068	0.096	0.041	0.08
G(Γ) × *Experiment*	-0.065	0.08	-0.082	0.119	-0.07	0.119	-0.074	0.114
G(Γ) × *Condition*	-0.029	0.038	0.006	0.044	-0.024	0.062	-0.002	0.055
S × *Frankfurt*	-0.015	0.074	-0.037	0.077	-0.01	0.071	-0.032	0.079
S × *Experiment*	0.092	0.084	0.097	0.107	0.085	0.082	0.089	0.109
S × *Condition*	0.042	0.048	0.056	0.05	0.04	0.047	0.055	0.052
Constant	0.169	0.205	-0.284*	0.144	0.116	0.179	-0.267*	0.143
R^2	0.587		0.398		0.585		0.429	
	F test for linear restrictions (all *n.s.*):							
	F(6, 89) = 1.24		F(6, 75) = 0.95		F(6, 89) = 1.62		F(6, 75) = 1.05	
	t	S.E.	t	S.E.	t	S.E.	t	S.E.
G(Γ)	0.374***	0.03	0.171***	0.033	0.383***	0.044	0.201***	0.037
S	0.524*	0.36	1.209***	0.278	0.638**	0.317	1.194***	0.269
Constant	0.165	0.195	-0.279*	0.143	0.112	0.174	-0.261*	0.142
R^2	0.573		0.375		0.574		0.407	

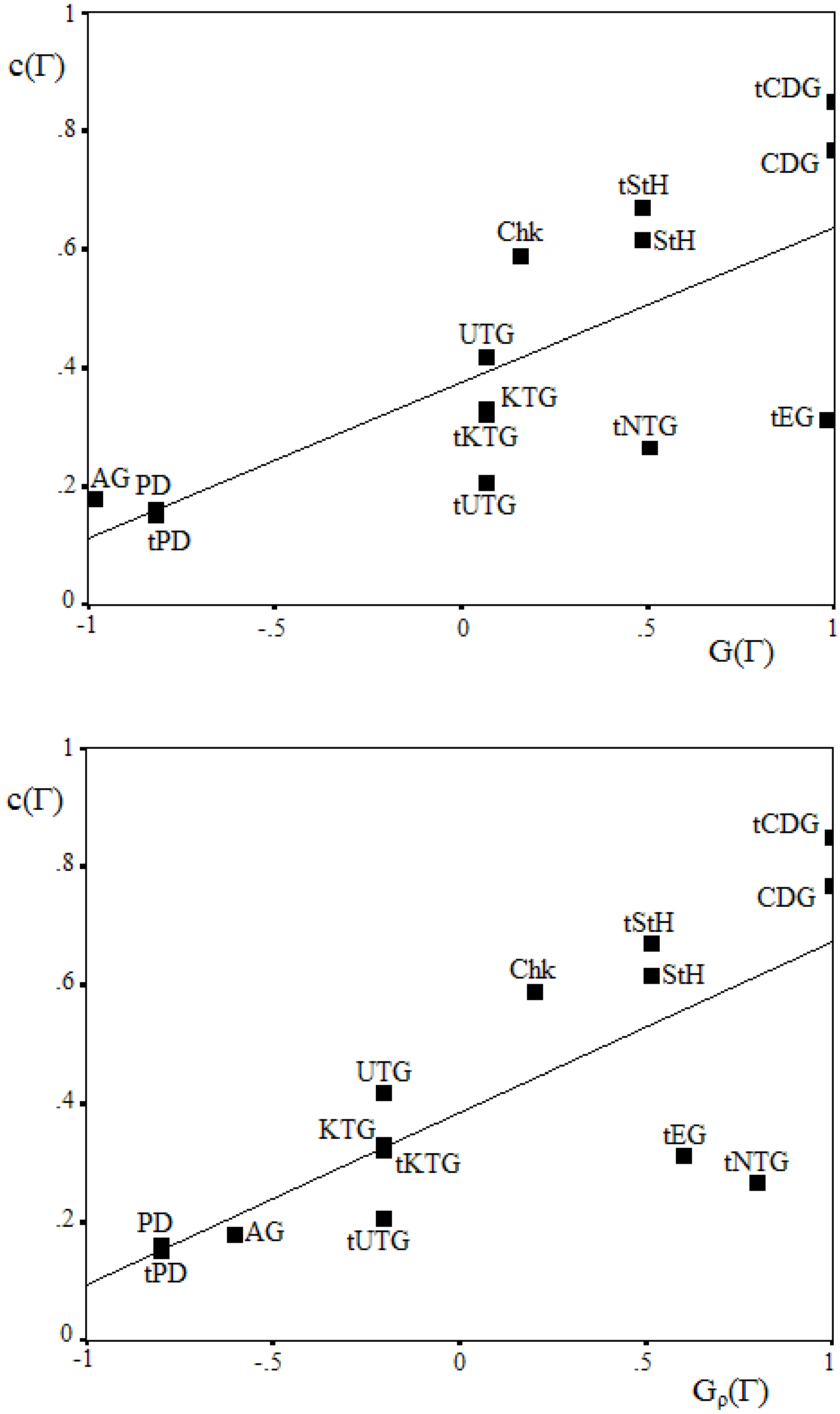
TABLE 7. PAIRS OF GAMES WHERE GAME HARMONY DOES NOT WORK

Prisoner's Dilemma		PD with Swapped Cells	
2, 2	0, 3	2, 2	0, 3
3, 0	1, 1	1, 1	3, 0

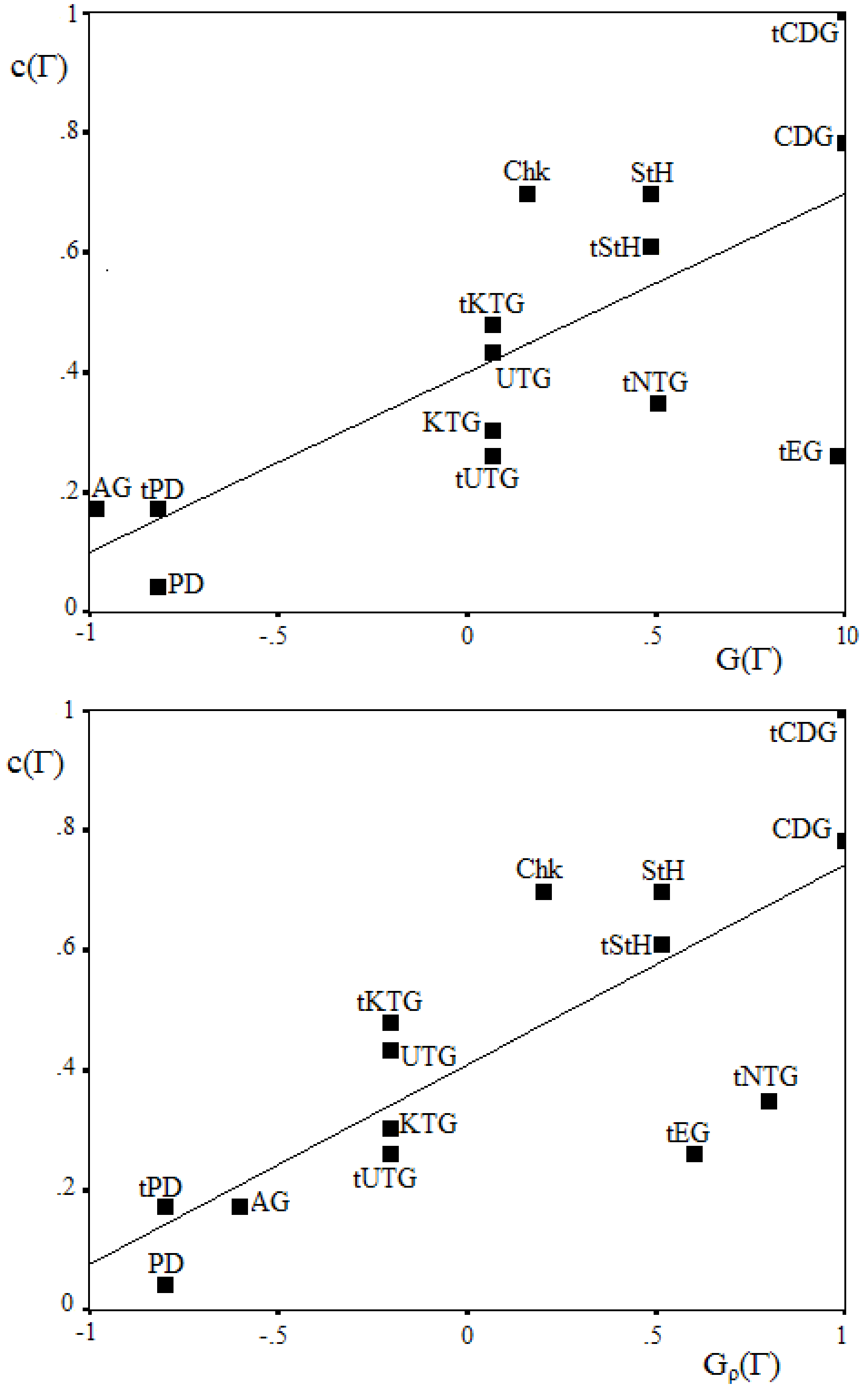
The game to the right is obtained by swapping the second row of the Prisoner's Dilemma (PD). The two games have identical $G(\Gamma)$ and $G_\rho(\Gamma)$ values, but in the game to the right there is no strictly dominant strategy or pure Nash equilibrium, suggesting that higher cooperative (*Top, Left*) play may be expected.

FIGURE 1. Basic Game Harmony Measures and R-Cooperation, Full Sample

Experiment 1



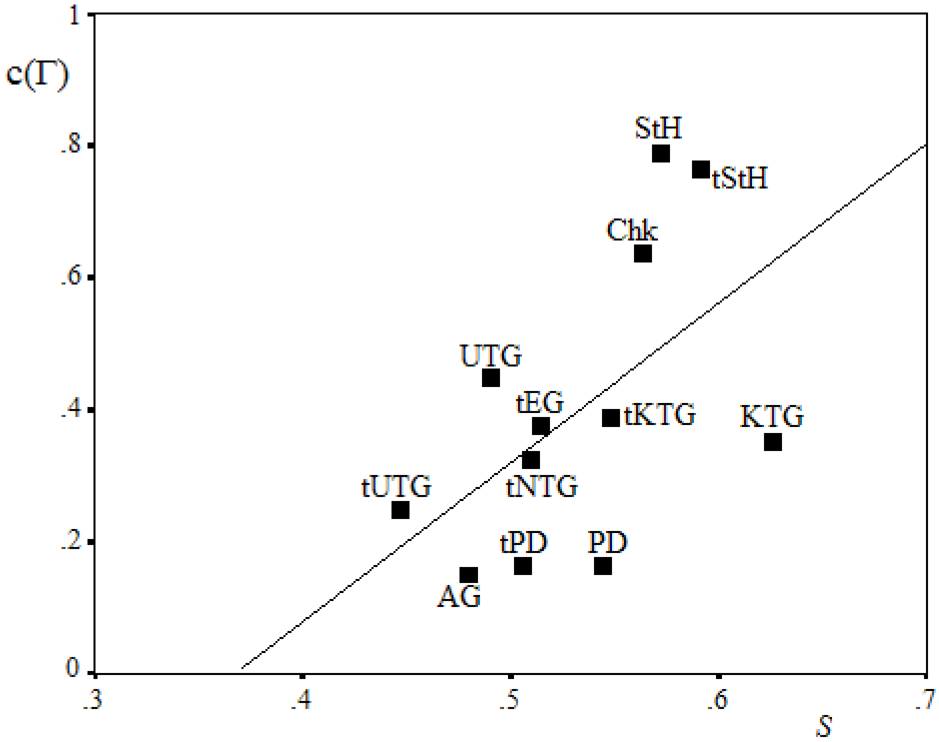
Experiment 2



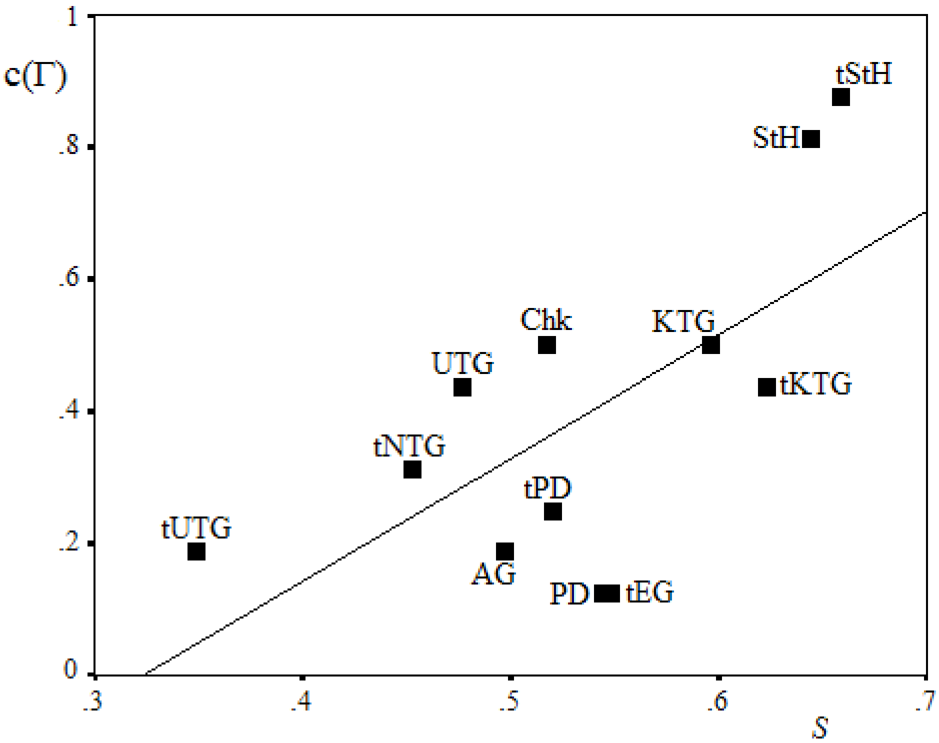
$c(\Gamma)$ is the mean r-cooperation rate. AG: Altruism G.; CDG: Coordination G.; ChK: Chicken; EG: Envy G.; PD: Prisoner's Dilemma; KTG/NTG/UTG: Kind/Needy/Unequitable Trust G., respectively; StH: Stag-Hunt. t*x* refers to the transposed presentation of game *x*.

FIGURE 2. Similarity Index and R-Cooperation, Restricted Sample

Experiment 1



Experiment 2



c(Γ) is the mean r-cooperation rate. AG: Altruism G.; CDG: Coordination G.; ChK: Chicken; EG: Envy G.; PD: Prisoner's Dilemma; KTG/NTG/UTG: Kind/Needy/Unequitable Trust G., respectively; StH: Stag-Hunt. t*x* refers to the transposed presentation of game *x*.