Joint Discussion Paper Series in Economics

by the Universities of

Aachen · Gießen · Göttingen
Kassel · Marburg · Siegen

ISSN 1867-3678

No. 27-2013

Hannes Rusch

What Niche did Human Cooperatives Evolve In?

This paper can be downloaded from
http://www.uni-marburg.de/fb02/makro/forschung/magkspapers/index_html%28magks%29

Coordination: Bernd Hayo • Philipps-University Marburg
Faculty of Business Administration and Economics • Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de

What niche did human cooperativeness evolve in?

Hannes Rusch[1-3]

Version: May 10, 2013

Abstract: The Prisoner's Dilemma (PD) is widely used to model social interaction between unrelated individuals in the study of the evolution of cooperative behaviour in humans and other species. Many effective mechanisms and promotive scenarios have been studied which allow for small founding groups of cooperative individuals to prevail even when all social interaction is characterised as a PD. Here, a brief critical discussion of the role of the PD as the most prominent tool in cooperation research is presented, followed by two new objections to such an exclusive focus on PD-based models of social interaction. It is highlighted that only 2 of the 726 combinatorially possible strategically unique ordinal 2x2 games have the detrimental characteristics of a PD and that the frequency of PD-type games in a space of games with random payoffs does not exceed about 3.5%. Although these purely mathematical considerations do not compellingly imply that the relevance of PDs is overestimated, it is proposed that, in the absence of convergent empirical information about the ancestral human social niche, this finding can be interpreted in favour of a so far rather neglected answer to the question of how the founding groups of human cooperation themselves came to cooperate: Behavioural and/or psychological mechanisms which evolved for other, possibly more frequent, social interaction situations might have been applied to PD-type dilemmas only later. Human cooperative behaviour might thus partly have begun as a cooptation.

Keywords: cooperation, prisoner's dilemma, cooptation, social niche, human evolution
JEL: C73; D74; D79

---

[1] JLU Giessen, Behavioral and Institutional Economics (VWL6), Licher Strasse 66, D-35394, Giessen, Germany, hannes.rusch@uni-giessen.de

[2] JLU Giessen, Philosophy of Biology Group, Associated member

[3] TU München, Peter Löscher Chair of Business Ethics, Arcisstrasse 21, D-80333 Munich

1. Why reconsider the structure of the human social niche?

Humans have been found to possess strikingly high levels of cooperativeness (Clutton-Brock 2009; Jaeggi et al. 2010; Jensen et al. 2007), raising challenging research questions for many disciplines, especially biology and economics (Hammerstein and Hagen 2005; Rankin 2011). However, the problem if and how cooperativeness can evolve under natural selection is not restricted to humans, of course. At the core of a majority of game theoretical models used in the study of the evolution of cooperativeness stands the (N-person) Prisoner's Dilemma ('PD'; Rapoport and Chammah 1965; Axelrod and Hamilton 1981; Boyd 1988). In situations of genuine PD-type all participating parties do best by always defecting (i.e. 'playing ALLD'). Their dominant strategies lead them directly into a Pareto-inefficient Nash-equilibrium. Many evolutionary trajectories have been proposed along which humans might have come to evolve unique psychological and/or cultural capacities for overcoming the PD, but none of them has proven fully satisfying (Fehr and Fischbacher 2003; Boyd and Richerson 2006; Nowak 2012). In line with many other authors (e.g. Noë 1990; Dugatkin 1992; West et al. 2007a; Tomasello 2009; Tomasello et al. 2012), I would like to encourage a different way of looking at the problem. In the following I present yet another argument for thoroughly reconsidering the assumption that the social niche which human cooperativeness evolved in is best characterised solely by PD-type interactions. Moving away from exclusively PD-based models of the evolution of human cooperativeness allows for other socio-ecological scenarios to be considered, which can be argued to be at least as instructive as the common PD-scenarios and may have greater potential to result in satisfying explanations of why humans became such an outstandingly cooperative species.

2. How cooperativeness can spread is well described, but how did it start?

When interactions are modelled as one-shot PDs played by unrelated, randomly matched individuals, permanent defection prevails if no additional mechanisms (i.e. means for strategy selection or payoff modification), like e.g. punishment (e.g. Clutton-Brock and Parker 1995), or partner choice (e.g. Johnstone and Bshary 2008), or additional assumptions e.g. about population structure (e.g. Ohtsuki et al. 2006), asymmetries (e.g. Dawkins 2010) or interdependencies (e.g. Roberts 2004) between players are introduced. Various effective mechanisms and promotive scenarios have been studied in theory which allow for groups of more cooperative individuals to invade and stabilise or even grow to dominate in such environments (see e.g. Nowak 2012). A common feature of many of these theories, partly owing to the use of local ESS analysis as the only mathematical tool, is that they need to assume the existence of at least some individuals ('founders') equipped with more elaborate strategy selection mechanisms which interact frequently enough to reap the benefits of cooperation and thereby gain relative fitness advantages over defectors.

But where do these founders of cooperation come from? Why did they evolve the more complicated and often slightly more vulnerable mechanisms they then apply in the otherwise defective environ-

ment? Now that we have gained a good overview of how cooperativeness can spread, this, I think, becomes the next challenging question.

Where theoretical works using the PD-paradigm do discuss the origins of the more cooperative strategies they investigate, it is sometimes assumed that simple one-step mutations can lead to the appearance of a sufficiently large number of 'hopeful monsters' (Binmore 2006) to start the transformation process in a population. For a number of reasons, which I will briefly outline in the next section (3), it is doubtful that simple one-step mutations in a PD environment are sufficient as explanations of the origins of cooperativeness (cf. e.g. André and Baumard 2011; Binmore 2006; Connor 1995; Dugatkin et al. 1992; Leimar and Hammerstein 2010; Skyrms 2004; Tomasello 2009; Tomasello et al. 2012). Some authors argue that the mechanisms that allow for cooperative solutions of PDs evolved for other problems and were applied to the PD only later (Binmore 2006; Tomasello 2009; Tomasello et al. 2012). After a brief review of the more general strengths and weaknesses of the PD-paradigm, I would like to try to make one more point in favour of this view.

3. Why study Prisoner's Dilemma models at all?

A simple answer to this question would be that only in genuine PD-type situations a cooperation problem arises. Thus, 'cooperation' here would be defined as 'playing C' in a PD and could therefore only be studied in PD-type interactions. This conceptual rigorousness would exclude phenomena like giving and helping (Bshary and Bergmüller 2008) or mutually beneficial coordinated behaviour (Mesterton-Gibbons and Dugatkin 1997) from the study of the evolution of cooperativeness, which, many think, is too restrictive—mostly since including these phenomena has proven very fruitful (West et al. 2007b; Leimar and Hammerstein 2010). A second answer is that the PD poses the "most stringent" cooperative dilemma (Nowak 2012) and thus can be used for 'stress testing' cooperation enabling mechanisms in theory, following the logic of 'if cooperation can make it there, it can make it anywhere.' I agree with this view—although, to my knowledge, the generalizability of results from PD models to other games has not been systematically investigated (Bshary and Bergmüller 2008; for advances in this direction see, e.g.: Dugatkin et al. 1992; Hauert et al. 2006; Kümmerli et al. 2007; Santos et al. 2011). Acknowledging the attractiveness of the PD as a performance benchmark in theoretical studies, we then have to ask, why we do not find more congruence of the hypothesised mechanisms with real animal behaviour (Noë 1990; Clutton-Brock 2009). There are at least two ways of answering this (which are not completely independent from one another).

First, one can review the, sometimes rather implicit, assumptions made by PD models about the situation in which the game actually is played and ask for each of these how realistic they are. They include: repeated or one-shot interaction, relatedness, population size and structure, anonymity and capabilities of partner recognition, memory constraints, 'trembling hands', availability of punishments/rewards, matching rules/partner choice, synchronous decision making, forced play, availability

of reputation information and its dynamics, symmetry, unavailability of third strategic options and more—see e.g. Raihani and Bshary (2011) and Nowak (2012) for discussions. The length of this listing already suggests that 'the pure PD situation' might be quite difficult to find in reality. Many, maybe too many, requirements have to be met by reality to match the theoretical models.

Second, it might simply be that no real world interaction actually has PD-type payoffs. There are reasons to doubt that (Raihani and Bshary 2011). Nevertheless, empirically it has turned out quite challenging to identify real world situations, for which the resulting 'net payoff' matrix of the game effectively played still has the features of a PD, when all relevant influences are taken into account. Thus, accepting that real social interaction situations exist which *prima facie* seem to possess PD-type payoffs, one could still argue that this impression very frequently disappears when looking more closely at the situation the game is embedded in. If, for example, interdependence between the players is allowed, the payoff matrix eventually changes to take the form of a 'Stag Hunt' game (SH; Roberts 2004). Also, if the PD game is repeated with the same partner with a certain probability, the net payoffs can be argued to take the form of a SH under some conditions (Skyrms 2004). These examples show that the two ways of criticising the PD are not independent: changing assumptions about the situation in which the game is played, frequently, if not always, changes the effective net payoffs as well.

Thus, this seems to be the heart of the debate about the use of PD models: Are there real situations in which the effective net payoffs actually form a PD? And if so: how frequent are they in nature, or, focussing on humans, how frequent and important were they in our ancestral past? I do not even want to start to try to answer this question empirically here. Instead, I would like to propose a new, admittedly rough, way of assessing the a priori relevance of PD-type games.

4. Can we assess the relevance of the Prisoner's Dilemma a priori?

For various reasons, including those just given, it is a very difficult task to determine the actual payoffs for the decision problems that ancestral populations (and species) were facing (see e.g. Alvard and Nolin 2002; Bird et al. 2012). If we knew these payoffs, speculation about the niche in which human cooperativeness evolved would be over. But since we do not, it might be worthwhile to try to simply examine all possible distributions of payoffs and see if that structure already suggests anything. Since most classical and a majority of recent theoretical works on the spread of cooperative traits do also do so, and for reasons of analytical feasibility, let us restrict a first analysis to interactions of two individuals facing binary decisions, i.e. classical 2x2 games. Subsequent analysis should of course also investigate games with more strategies (like e.g. Worden and Levin 2007) and/or more players (see e.g. Connor 2010).

The infinite space of possible payoff matrices can be reduced to a manageable size by restricting analyses to non-strictly ordered preference distributions: i.e. each player can have up to four different pref-

erences for interaction outcomes, indicated by 1 to 4, 4 being the most preferred outcome. The finite but still ample space of games spanned by all combinations of possible preference distributions can then again be reduced by eliminating strategically equivalent games. Two games are strategically equivalent if one can be obtained from the other by renaming players and/or their respective strategies. At the end of this computation 726 unique matrices remain which can be studied by exhaustion (see Appendix 1; Guyer and Hamburger 1968; Fraser and Kilgour 1986). Recently, a very elegant way of topologically arraying preference distributions has been put forward (Robinson and Goforth 2005). That method focuses only on strictly ordered preferences and does not exclude strategically equivalent distributions. However, the following observations mutatis mutandis also hold for that subclass of distributions, of course.

The crucial feature of PD-type situations is that both players have strictly dominant strategies which in combination lead to a Pareto-inefficient outcome. Among the 726 possible games, there are 91 in which both players possess a dominant strategy. But only two of these are PD-type (Table 1): the PD, of course, and an asymmetrical variant in which one of the players cannot benefit from mutual cooperation while the other could.

	C	**D**
C	3 : 3	1 : 4
D	4 : 1	2 : 2

Prisoners' Dilemma (PD)
Single inefficient Nash-equilibrium (D,D)

	C	**D**
C	2 : 3	1 : 4
D	3 : 1	2 : 2

Asymmetrical variant of PD
Single inefficient Nash-equilibrium (D,D)

Table 1: The two PD-type situations

Thus, only 2 of 726 (=0.28%) preference distributions have the detrimental characteristics of a PD. In contrast to that, the combination of both players' dominant strategies leads to an at least Pareto-efficient outcome in 89 of 726 cases (=12.26%). In 40 of these 89 (=5.5%) the outcome is optimal for one of the players. In 36 cases (=4.96%) it is even optimal for both. A subset of the remaining 635 (=87.47%) preference distributions are those 234 cases (=32.23%) in which one player has a dominant strategy and the other has a best response to that strategy. When the later chooses that best response, the resulting outcome is at least Pareto-efficient in 214 cases (=29.48%). It is optimal for one player in 108 cases (=14.88%) and optimal for both in 96 cases (=13.22%). However, the remaining 20 cases (=2.75%) in this subset have the character of a milder dilemma. In 8 of these cases (=1.10%) the Pareto-inefficient result is optimal for one player. Thus, Pareto-inefficiency here implies that this player would have to be brought to switch strategies to allow for the improvement of the other player's outcome while potentially risking her own optimal outcome. In the remaining 12 cases (=1.65%) the Pareto-inefficient outcome is optimal for none of the players. In 6 (=0.83%) of these cases, switching to an alternative strategy by one or both players could at best benefit one of them. In the remaining 6

cases both could benefit from switching strategies simultaneously. These six cases are the so called 'Alibi games' which are the closest relatives of the PD (Robinson and Goforth 2005).

In the remaining 401 (=55.23%) conceivable games the concept of dominant strategies cannot be applied. For that subset other means for strategy selection have to be investigated, e.g. risk-dominance analysis. A broader systematic analysis of the whole space of distributions promises further insights but would go beyond the scope of the argument here (Hauert et al. 2006).

Strategic scenario	**Count (%)**
No dominant strategies	401 (55.23)
One dominant strategy, Pareto-efficient outcome	214 (29.48)
Two dominant strategies, Pareto-efficient outcome	89 (12.26)
Mild dilemmas: one dominant strategy, Pareto-inefficient outcome, optimal for one of the actors	8 (1.10)
Mild dilemmas: one dominant strategy, Pareto-inefficient outcome, optimal for none of the actors, switching would only benefit one actor	6 (0.83)
Alibi games: one dominant strategy, Pareto-inefficient outcome, optimal for none of the actors, switching could benefit both actors	6 (0.83)
PD-type games	2 (0.28)

Table 2: Summary of game counts for various strategic situations

Although the figures stated (summarised in Table 2) are derived from very basic combinatorial and game theoretical considerations, I believe they suffice to show that PD-type situations are outliers in a structural sense. In order to assess the ecological relevance of PD-type problems more thorough quantitative investigations of our most likely ancestral (social) ecology need to be made, of course (as do e.g. Hawkes 1993; Hill 2002). What I have tried to show here is merely that from an a priori mathematical perspective, in almost all conceivable 2x2 social interactions no PD-type situation arises.

At a closer look, this result has a trivial appeal. And indeed, instead of enumerating all 726 games and counting those which match the PD criterion, we can also take a shortcut. Using the standard notation for 2x2 games (payoffs $X \in \{T, R, P, S\}$) but allowing for asymmetry (payoffs X_i with $i = 1,2$ denoting the respective player) the PD definition demands that both players have a strictly dominant strategy, yielding (w.l.o.g.) $T_i > R_i$ and $P_i > S_i$, and that the combination of these strategies must be Pareto-inefficient, thus (i) $P_i < R_i$ or (iia) $P_1 < R_1$ and $P_2 = R_2$ or (iib) $P_1 = R_1$ and $P_2 < R_2$. Conditions (iia) and (iib) are obtainable from one another by renaming players and thus yield only one strategically unique game which satisfies them. Translated into preferences this directly results in the two games given in Table 1. Thus, in the space of the 726 strategically unique 2x2 games in preferences PD-type interactions can be said to be rare by definition.

The next step, and maybe the last one that can be taken a priori, is thus to move away from the 726 games set and try to estimate the frequency with which PD-type interactions occur when payoffs are randomly chosen. Since it is a combinatorially rather cumbersome enterprise to calculate the respective probabilities with which different classes of payoff matrices occur in such a process, I only present the results of a numerical simulation. Figure 1 shows the mean frequencies f_n with which PD-type games arise from the following process: eight integers are drawn with replacement from a uniform distribution of the integers 1 through n. The integers are interpreted as a 2x2 game represented by the vector (T_1, R_1, P_1, S_1, T_2, R_2, P_2, S_2) and it is checked whether this game or one of its seven equivalent representations matches the PD definition as stated above. This is repeated 100.000 times for every n. Starting at $f_4 \approx 0.007$ the observed mean PD frequencies converge to a value $f_\infty \approx 0.034$. Thus, interestingly, when randomly creating 2x2 games, the mean observed frequency of PD-type games is more than ten times higher than in the 726 games set but still quite small.

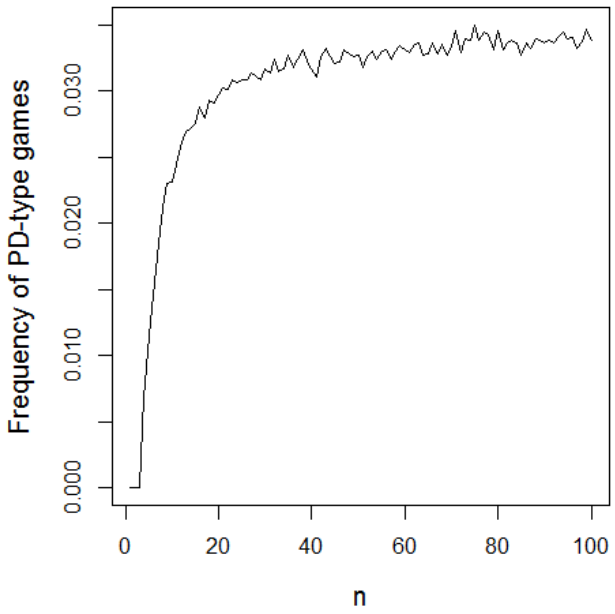

Figure 1: Observed mean frequencies of PD-type games when randomly creating 2x2 games with payoffs drawn with replacement from a uniform distribution of integers from 1 to n

The interesting question is now: In what way—if at all—do these small a priori frequencies of PD-type games relate to real social interactions? It could well be that the costs and benefits of real interactions are somehow constrained in a way that makes PD-type payoff structures much more likely. A thorough investigation of this question seems overdue, but would encompass substantial empirical work and go beyond the scope of the reasoning presented here—for a related discussion see Raihani and Bshary (2011).

To make this point most explicit: Both a priori considerations presented here, analytical and by simulation, do by no means show that the PD is ecologically irrelevant; but even the space of 2x2 games,

obviously, has a lot more to offer. This becomes relevant e.g. when we remember that small changes in the assumptions about the real situation which games are embedded in can change the effective net payoffs to form other games. The study of these other games, in my view, is a key to understanding where and why mechanisms for strategy choice evolved which also 'do the trick' in PDs.

5. A cooptation hypothesis regarding cooperative behaviour

Non-PD situations are not at all unproblematic, no doubt (Hauert et al. 2006; Kümmerli et al. 2007). Many of them represent intricate decision problems, like the well-known 'Stag Hunt' (Skyrms 2004) or the 'Hawk-Dove' game (Kümmerli et al. 2007; Smith and Price 1973). Nonetheless, in environments dominated by problems of non-PD-type it is possible to conceive of evolutionary trajectories toward ever greater cooperativeness without having to make use of 'hopeful monsters'. Step-by-step natural selection can lead toward cooperation for mutual benefits in environments like these much easier (Binmore 2006; Leimar and Hammerstein 2010; Tomasello 2009; Tomasello et al. 2012). One scenario suggests itself quickly here: the transfer of behavioural mechanisms which are adaptive for interaction with kin to interactions with non-kin (see e.g. Axelrod and Hamilton 1981), an idea that has been called a 'big mistake hypothesis' (Boyd and Richerson 2006; Tomasello et al. 2012). Although I do think that this idea deserves more attention than Boyd and Richerson (2006) are willing to give to it, a recent empirical finding suggests that our ancestral social groups might have been composed of less kin than thought so far (Hill et al. 2011). This finding emphasises that we should also try to analyse ways by which cooperation among non-kin could have evolved from scratch. Prominent examples of cooperative bonds between unrelated individuals, which call for explanations that probably cannot be based on genetic relatedness at all, are interspecific mutualisms, like e.g. in certain plants and their fungal symbionts (e.g. Kiers et al. 2011) or in cleaner fish (*Labroides dimidiatus*; Bshary and Grutter 2003). But also cooperative behaviour in primates might frequently be better interpreted as solving coordination problems in non-PD-type situations than as solving Prisoner's Dilemmas (Noë 1990). Chimpanzees (*Pan troglodytes;* Bullinger et al. 2011), rhesus monkeys (*Macaca mulatta*) and capuchin monkeys (*Cebus apella*; Brosnan et al. 2012) e.g. have been found to readily behave cooperatively in mutually beneficial Stag Hunt games.

The cooptation hypothesis—see Buss et al. (1998) for a disambiguation of the term 'cooptation'—now proposes that, once evolved, mechanisms for coordination on mutually beneficial outcomes in non-PD situations, like e.g. non-binding communication (Brosnan et al. 2012; Koukoumelis et al. 2012; Noë 2006; Ostrom et al. 1992) or unilateral passive leadership (Bullinger et el. 2011), can also be successfully applied in PD-type situations. It would be worth investigating if this holds, e.g., for the basic partner control and partner choice mechanisms in repeated interaction. The well-known GRIM mechanism ('completely stop cooperating once you have been defected against') for example does not only beat ALLD in iterated PDs under certain circumstances but can also lead to coordination on the payoff

dominant equilibrium in iterated SHs (because GRIM actually makes every strategy combination that has a higher payoff than its achievable alternatives a Nash-equilibrium in games which are repeated sufficiently long enough, presumed that the other players are also playing GRIM; see Rubinstein 1979). The crucial point of the example is: if GRIM is thought of as a mechanism evolved for solving PDs, we would have to explain how a friendly strategy (GRIM cooperates until it is defected against) could have gained foothold in an otherwise defective population where it would have been exploited perpetually (i.e. in the 'first round' of every encounter). If we assume instead that GRIM evolved as a mechanism for stabilising mutually beneficial coordination in SH-like situations and was applied to severer dilemmas only later, then this explanatory problem simply disappears.

Adding to this picture, it has been shown recently, to name a last example, that pre-play signalling, a mechanism which also readily evolves for coordination on mutually beneficial outcomes in Stag Hunt games, can destabilise the full defection equilibrium in PD games (Santos et al. 2011).

I hope these first examples suffice to back up the idea that mechanisms allowing for cooperative solutions of PD-type situations might partly be cooptations of mechanisms which evolved for other, maybe more frequent, coordination problems.

6. Conclusion

Social interactions of a kind in which it is not always the best option to defect, I think, form an indispensable bridge between kin selection driven evolution and the high levels of cooperativeness among unrelated strangers we observe in humans today. It is easier to understand why we seem to possess such distinctive inclinations to cooperate with conspecifics, when our ancestors thrived by successfully solving more and more coordination problems to mutual benefit, than when we conceive of them facing Prisoner's Dilemmas every time they interacted (Tomasello et al. 2012). The contrast between the multitude of theoretical works on the evolution of cooperation using PD-type model situations and the sparseness of empirical evidence for the ecological validity of their findings in animals (Noë 2006; Clutton-Brock 2009) can be seen as a hint that some important link is missing here. In line with other authors, I would like to emphasise that we need to broaden the scope of search for mechanisms enabling cooperative behaviour also in humans. We should include mechanisms that offer solutions to milder coordination problems and investigate what happens when these are applied to problems of severer dilemma types.

In short: From the structural perspective outlined here, it seems plausible that our evolutionary ancestors discovered more and more ways to reap mutual benefits and finally also applied the psychological and/or behavioural mechanisms they evolved for doing so to problems of PD-type. Thus, the social niche (human) cooperativeness evolved in, I think, is best conceived of as a colourful mix of many coordination problems and only some severe dilemmas.

Appendix 1: The 726 strategically unique 2x2 games in preference distributions

A 2x2 game has two players ('R(ow)' and 'C(olumn)') who each have two strategies ('U(p)'/'D(own)' for R and 'L(eft)'/'R(ight)' for C). This yields four unique strategy combinations: UL, DL, UR, and DR. Each player can now assign up to four different preference values to these strategy combinations, whereby higher values indicate stronger preference. If, e.g., one player had no preferences at all, she would assign the value 1 to all four strategy combinations. If, e.g. R preferred UL and UR over DL and DR, but did not care whether L or R was played by C, she would assign two different preferences like this: UL~2, UR~2, DL~1, DR~1. (Thus, in this example, R does not have to take into account what C chooses, since C's choice cannot change the outcome for R.) Every 2x2 game given in absolute payoffs can be mapped onto the respective game in preference distributions simply by replacing every payoff with its rank in the set of all four payoffs for that player: the payoff set (13, 29, 61, 19), e.g., would thus be transformed to the preference distribution (1, 3, 4, 2). Note that this mapping function is surjective but non-injective: one game in preference distributions corresponds to an (infinite) set of games in absolute payoffs.

There are 726 strategically unique 2x2 games when the four possible strategy combinations are given as non-strictly ordered preference distributions for each of the two players. These 726 unique games can be obtained by simply enumerating all combinatorially possible games ($= 4^8 = 65\,536$ games) and including only those in the list of unique games, which are not strategically equivalent to a game which is already on the list. (A simple computer program for this purpose as well as the list of unique games with their strategic properties is available from the author upon request; see the original papers by Guyer and Hamburger (1968) and Fraser and Kilgour (1986) for additional details.)

Acknowledgements: I thank Max Albert and Eckart Voland for helpful criticism.

References

Alvard MS, Nolin DA (2002) Rousseau's whale hunt? Curr Anthropol 43:533-559.
André JB, Baumard N (2011) Social opportunities and the evolution of fairness. J Theor Biol 289:128-135.
Axelrod R, Hamilton WD (1981) The evolution of cooperation. Science 211:1390-1396.
Binmore K (2006) Why do people cooperate? Polit Philos Econ 5:81-96.
Bird RB, Scelza B, Bird DW, Smith EA (2012) The hierarchy of virtue: mutualism, altruism and signaling in Martu women's cooperative hunting. Evol Hum Behav 33:64-78.
Boyd R (1988) Is the repeated prisoner's dilemma a good model of reciprocal altruism? Ethol Sociobiol 9:211-222.
Boyd R, Richerson PJ (2006) Solving the puzzle of human cooperation. In: Levinson SC (ed) Evolution and Culture. MIT Press, Cambridge, pp 105-132
Brosnan SF, Wilson BJ, Beran MJ (2012) Old World monkeys are more similar to humans than New World monkeys when playing a coordination game. Proc R Soc B 279:1522-1530.
Bshary R, Bergmüller R (2008) Distinguishing four fundamental approaches to the evolution of helping. J Evol Biol 21:405-420.
Bshary R, Grutter AS (2003) Asymmetric cheating opportunities and partner control in a cleaner fish mutualism. Anim Behav 63:547-555.
Bullinger AF, Wyman E, Melis AP, Tomasello M (2011) Coordination of chimpanzees (Pan troglodytes) in a Stag Hunt game. Int J Primatol 32:1296-1310.
Buss DM, Haselton MG, Shackelford TK, Bleske AL, Wakefield JC (1998) Adaptations, exaptations, and spandrels. Am Psychol 53:533-548.
Clutton-Brock TH (2009) Cooperation between non-kin in animal societies. Nature 462:51-57.
Clutton-Brock TH, Parker GA (1995) Punishment in animal societies. Nature 373:209-216.
Connor RC (1995) Altruism among non-relatives: alternatives to the 'Prisoner's Dilemma'. Trends Ecol Evol 10:84-86.
Connor RC (2010) Cooperation beyond the dyad: on simple models and a complex society. Proc R Soc B 365:2687-2697.
Dawkins MS (2010) Do asymmetries destabilize the Prisoner's Dilemma and make reciprocal altruism unlikely? Anim Behav 80:339-341.
Dugatkin LA, Mesterton-Gibbons M, Houston AI (1992) Beyond the prisoner's dilemma: Toward models to discriminate among mechanisms of cooperation in nature. Trends Ecol Evol 7:202-205.
Fehr E, Fischbacher U (2003) The nature of human altruism. Nature 425:785-791.
Fraser NM, Kilgour DM (1986) Non-strict ordinal 2x2 games. Theory Decis 20:99-121.
Guyer M, Hamburger H (1968) A note on "A taxonomy of 2x2 games". Gen Syst 13:205-208.
Hammerstein P, Hagen EH (2005) The second wave of evolutionary economics in biology. Trends Ecol Evol 20:604-609.
Hauert C, Michor F, Nowak MA, Doebeli M (2006) Synergy and discounting of cooperation in social dilemmas. J Theor Biol 239:195-202.
Hawkes K (1993) Why hunter-gatherers work. Curr Anthropol 34:341-361.
Hill KR (2002) Altruistic cooperation during foraging by the Ache and the evolved human predisposition to cooperate. Hum Nat 13:105-128.
Hill KR, Walker RS, Božičević M, Eder J, Headland T, Hewlett B, Hurtado AM, Marlowe F, Wiessner P, Wood B (2011) Co-residence patterns in hunter-gatherer societies show unique human social structure. Science 331:1286-1289.
Jaeggi AV, Burkart JM, van Schaik CP (2010) On the psychology of cooperation in humans and other primates: combining the natural history and experimental evidence of prosociality. Philos Trans R Soc B 365:2723-2735.
Jensen K, Call J, Tomasello M (2007) Chimpanzees are rational maximizers in an ultimatum game. Science 318:107-109.

Johnstone RA and Bshary R (2008) Mutualism, market effects and partner control. J Evol Biol 21:879-888.
Kiers ET, Duhamel M, Beesetty Y, Mensah JA, Franken O, Verbruggen E, Fellbaum CR, Kowalchuk GA, Hart MM, Bago A, Palmer TM, West SA, Vandenkoornhuyse P, Jansa J, Bücking H (2011) Reciprocal rewards stabilize cooperation in the mycorrhizal symbiosis. Science 333:880-882.
Koukoumelis A, Levati MV, Weisser J (2012) Leading by words: A voluntary contribution experiment with one-way communication. J Econ Behav Organ 81:379-390.
Kümmerli R, Colliard C, Fiechter N, Petitpierre B, Russier F, Keller L (2007) Human cooperation in social dilemmas: comparing the Snowdrift game with the Prisoner's Dilemma. Proc R Soc B 274:2965-2970.
Leimar O, Hammerstein P (2010) Cooperation for direct fitness benefits. Philos Trans R Soc B 365:2619-2626.
Mesterton-Gibbons M, Dugatkin LA (1997) Cooperation and the Prisoner's Dilemma: toward testable models of mutualism versus reciprocity. Anim Behav 54:551-557.
Noë R (1990) A Veto game played by baboons: a challenge to the use of the Prisoner's Dilemma as a paradigm for reciprocity and cooperation. Anim Behav 39:78-90.
Noë R (2006) Cooperation experiments: coordination through communication versus acting apart together. Anim Behav 71:1-18.
Nowak MA (2012) Evolving cooperation. J Theor Biol 299:1-8.
Ohtsuki H, Hauert C, Lieberman E, Nowak MA (2006) A simple rule for the evolution of cooperation on graphs and social networks. Nature 441:502-505.
Ostrom E, Walker J, Gardner R (1992) Convenants with and without a sword. Am Polit Sci Rev, 86:404-417.
Raihani NJ, Bshary R (2011) Resolving the iterated prisoner's dilemma: theory and reality. J Evol Biol 24:1628-1639.
Rankin DJ (2011) The social side of Homo economicus. Trends Ecol Evol 26:1-3.
Rapoport A, Chammah AM (1965) Prisoner's dilemma. University of Michigan Press, Ann Arbor
Roberts G (2004) Cooperation through interdependence. Anim Behav 70:901-908.
Robinson D, Goforth D (2005) The topology of the 2x2 games. Routledge, London
Rubinstein A (1979) Equilibrium in supergames with the overtaking criterion. J Econ Theory 21:1-9.
Santos FC, Pacheco JM, Skyrms B (2011) Co-evolution of pre-play signaling and cooperation. J Theor Biol 274:30-35.
Skyrms B (2004) The stag hunt and the evolution of social structure. Cambridge University Press, Cambridge
Smith JM, Price GR (1973) The logic of animal conflict. Nature 246:15-18.
Tomasello M (2009) Why we cooperate. MIT Press, Cambridge
Tomasello M, Melis AP, Tennie C, Wyman E, Herrmann E (2012) Two key steps in the evolution of human cooperation: the interdependence hypothesis. Curr Anthropol 53:673-692.
West SA, Griffin AS, Gardner A (2007a) Evolutionary Explanations for Cooperation. Curr Biol 17:R661-R672.
West SA, Griffin AS, Gardner A (2007b) Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. J Evol Biol 20:415-432.
Worden L, Levin SA (2007) Evolutionary escape from the prisoner's dilemma. J Theor Biol 245:411-422.