Hans-Ulrich Stark, Dirk Helbing, Martin Schönhof, and Janusz A. Hołyst:
Alternating cooperation strategies in a Route Choice Game:
Theory, experiments, and effects of a learning scenario.
In: A. Innocenti and P. Sbriglia (eds.): *Games, Rationality and Behaviour.*
Houndmills and New York: Palgrave McMillan, pp. 256–273, 2008.

Alternating cooperation strategies in a Route Choice Game: Theory, experiments, and effects of a learning scenario

Hans-Ulrich Stark[1], Dirk Helbing[2], Martin Schönhof[2], and Janusz A. Hołyst[3]

[1] *Chair of Systems Design, ETH Zurich, CH-8032 Zurich, Switzerland,* hustark@ethz.ch

[2] *Institute for Transport & Economics, Dresden University of Technology, Andreas Schubert Str. 23, 01062 Dresden, Germany*

[3] *Faculty of Physics and Center of Excellence for Complex Systems Research, Warsaw University of Technology, Koszykowa 75, PL-00-662 Warsaw, Poland*

Abstract

In this paper, we present experimental investigations on a day-to-day route choice scenario. Here, the equilibrium outcome is, according to real traffic observations, fair (equal for all users) but induces an inefficient usage of network capacity. Optimal usage would be characterized by some users winning and some losing in comparison to the equilibrium state. Coherent, alternating cooperation strategies can be a suitable solution, but they require innovation and group coordination in addition to cooperativeness. In these points, our work differs considerably from other contributions observing the emergence of cooperation in social dilemmas. By classifying the two-person variant of our experiments among the symmetrical 2x2 games we show the situation of the *Route Choice Game* not to be addressed by the literature so far. Although the equilibrium outcome in this setup is "strongly stable", in our experiments we find eminent empirical evidence of alternating cooperation and, thereby, observed persistent utilization of the system optimum. The transition to this optimal configuration can be well described by quantitative considerations that are presented in the paper. Furthermore, the straight success of a learning scenario indicates that the collective innovation of alternating strategies may be the most critical challenge to the individuals instead of just learning to be cooperative. Presumably, this also holds for other social dilemma situations.

Keywords: Game theory, experiments, alternating cooperation, learning

1 Introduction

The question of optimally distributed entity-flows in capacity-restricted networks is a certain kind of social dilemma (compare to [14, 16, 25, 34, 35, 43]). Inefficiencies in urban traffic networks, especially during peak times in the morning and afternoon, are not only inconvenient but also expensive, both from the individual and general public point of view. Nevertheless, the concept of the "Wardrop equilibrium"[1][48] is not only a theoretical construct, but can approximately also be observed in reality. It is described by equal travel times on all suitable routes between a given origin and destination ("o-d pair") which obviously does not necessarily imply a minimum of overall travel times. However, a system optimal distribution, characterized by minimal overall travel times, would require that a fraction of users additionally change to another alternative, thereby decreasing travel times on their previous route but increasing their own travel times on the new route[2].

Since Pigou [37], it has been suggested to resolve the problem of inefficient road usage by congestion charges, but are they really needed? Is the missing establishment of a system optimum just a problem of varying traffic conditions and changing origin-destination pairs, which make route-choice decisions comparable to one-shot games? Or would individuals in an *iterated* setting of a day-to-day route choice game with identical conditions spontaneously establish cooperation in order to increase their returns, as the folk theorem suggests [6]?

How would such a cooperation look like? Taking turns could be a suitable solution [42]. While simple symmetrical cooperation is typically found for the repeated Prisoner's Dilemma [3, 4, 28, 29, 30, 33, 36, 38, 41, 43, 46, 47], emergent alternating reciprocity has recently been discovered for the games Leader and Battle of the Sexes [9]. Note that such coherent oscillations are a time-dependent, but deterministic form of individual decision behavior, which can establish a persistent phase-coordination, while mixed strategies, i.e. statistically varying decisions, can establish cooperation only by chance or in the statistical average. This difference is particularly important when the number of interacting persons is small, as in the particular route choice game discussed below.

Note that oscillatory behavior has been found in iterated games before:

- In the rock-paper-scissors game [46], cycles are predicted by the game-dynamical equations due to unstable stationary solutions [19].
- Oscillations can also result by coordination problems [2, 20, 21, 22], at the cost of reduced system performance.
- Moreover, blinker strategies may survive in repeated games played by a mixture of finite automata [5] or result through evolutionary strategies [9, 10, 11, 23, 24, 26, 27, 49].

[1] Also referred to as *system optimum* or *user optimum*.
[2] The basic assumption here is that travel times at least on average monotonously increase with road occupancy.

However, these oscillation-generating mechanisms are clearly to be distinguished from the establishment of phase-coordinated, alternating reciprocity we are interested in (coherent, oscillatory cooperation to reach the system optimum).

In this paper, we review former results by the authors [18] but focus more on game theoretical implications and on the results of a learning scenario. The paper is organized as follows: In section 2, we introduce our empirical work on this topic. It is based on previous investigations by Selten *et al.* and Helbing *et al.* [17, 44] but extends their work by a crucial feature - the distinction between equilibrium and overall efficient network usage. This directly leads us to the question of the corresponding situation (game) in game theory. Section 3 provides the corresponding discussion and findings, thereby introducing the Route Choice Game as a symmetrical 2x2 game that is almost completely neglected in the literature so far. Additionally, at the end of section 2, quantitative investigations regarding the transition to coordinated alternating cooperation are presented. The final section 4 summarizes the obtained results and gives an outlook on promising fields of future research.

2 Experimental setup and results

2.1 Setup

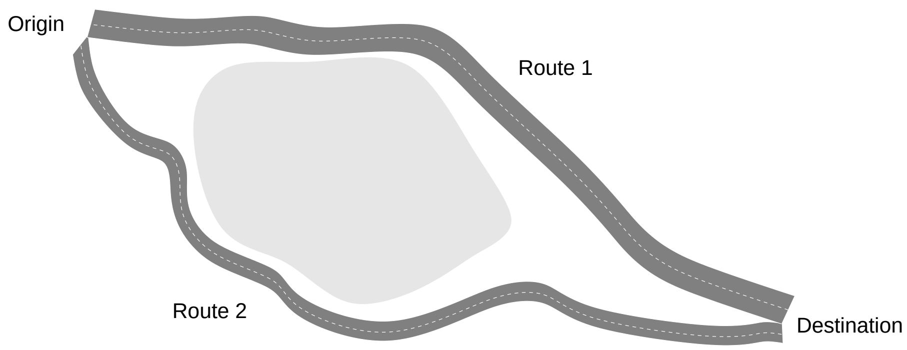

Figure 1: Illustration of the investigated day-to-day route choice scenario. We study the dynamic decision behavior in a repeated route choice game, where a given destination can be reached from a given origin via two different routes, a freeway (route 1) and a side road (route 2).

In the following, we will investigate a scenario with two alternative routes between a certain origin and a given destination, say, between two places or towns A and B (see Fig. 1). We are interested in the case where both routes have different capacities, say a freeway and a subordinate or side road. While the freeway is faster when it is empty, it may be reasonable to use the side road when the freeway is congested.

N_1	0	1	2
N_2	2	1	0
$P_1(N_1)$	-	300	0
$P_2(N_2)$	-200	-100	-

N_1	0	1	2	3	4
N_2	4	3	2	1	0
$P_1(N_1)$	-	600	300	0	-300
$P_2(N_2)$	-300	-200	-100	0	-

Table 1: Payoff tables for experiments with N=2 (left) and N=4 participants (right).

The "success" of taking route i could be measured in terms of its inverse travel time $1/T_i(N_i) = V_i(N_i)/L_i$, where L_i is the length of route i and $V_i(N_i)$ the average velocity when N_i of the N drivers have selected route i. One may roughly approximate the average vehicle speed V_i on route i by Greenshield's linear velocity-density relationship [15]

$$V_i(N_i) = V_i^0 \left(1 - \frac{N_i(t)}{N_i^{\max}}\right) , \tag{1}$$

where V_i^0 denotes the maximum velocity (speed limit) and $N_i^{\max}$ the capacity, i.e. the maximum possible number of vehicles on route i. With $A_i = V_i^0/L_i$ and $B_i = V_i^0/(N_i^{\max} L_i)$, the inverse travel time then obeys the relationship

$$1/T(N_i) = A_i - B_i N_i \, , \tag{2}$$

which is linearly decreasing with the road occupancy N_i. Other monotonously falling relationships $V_i(N_i)$ would make the expression for the inverse travel times non-linear, but they would probably not lead to qualitatively different conclusions.

Using this motivation, we define the payoff functions for our experiments.

$$P_i(N_i) = C_i - D_i N_i \tag{3}$$

Since there are only two routes, leading to a total of $N = N_1 + N_2$, we can say that the payoff of route 1 falls and the payoff of route 2 rises with higher occupation of route 1. We scaled the functions in a way that there is an integer fraction of route 1 choosers (N_1) that leads to equal payoffs for both routes (representing the Wardrop equilibrium). Furthermore, for a better orientation to the participants, this payoff value is set to zero. For the 2-participants setup we chose $C_1 = 600$, $D_1 = 300$, $C_2 = 0$, and $D_2 = 100$. For the 4-participants experiments we change C_1 to the value of 900 which leads to the payoffs shown in Table 1.

Altogether we have carried out more than 80 route choice experiments with different experimental setups, all with different participants. In the 24 two-person [12 four-person] experiments evaluated here, test persons were instructed to choose between two possible routes between the same origin and destination. They knew that route 1 corresponds to a 'freeway' (which may be fast or congested), while route 2 represents an alternative route. Test persons were also informed that, if two [three] participants chose route 1, everyone would receive 0 points, while if half of the participants chose route 1, they would receive the maximum average amount of 100 points per person, but 1-choosers would profit at the cost of 2-choosers. Finally, participants were told that everyone could reach an average of 100 points per round and that the (additional) individual payment after the experiment would depend on their cumulative payoff points reached in at least 300 rounds (100 points = 0.01 EUR). Of course, communication between the participants was not allowed. See the appendix for an example of the handout to the participants.

In the main experimental setup, there first were two separate two-person experiments simultaneously. Thereafter, all participants were merged into one group and played together in four-person experiments. This is what we refer to as a "learning scenario" and what will be further discussed and analyzed in section 2.4. This setup allowed us, on one hand, to investigate the participants' unaffected behavior in the iterated 2x2 Route Choice Game[3] (section 2.2) and, on the other hand, to explore learning effects on the behavior in four-person experiments.

2.2 Results: Emergence of cooperation

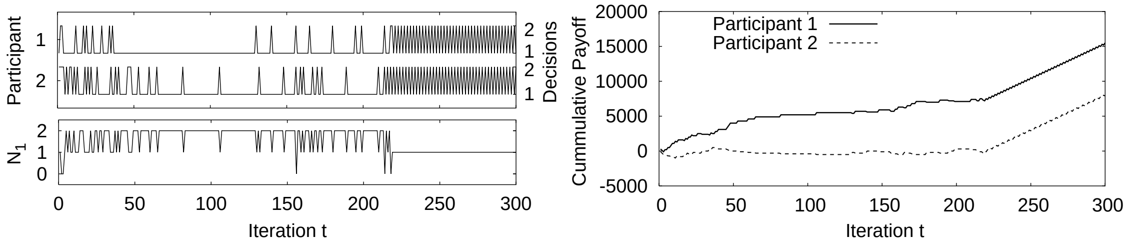

Figure 2: Representative example for the emergence of coherent oscillations in a 2-person route choice experiment with the parameters specified in table 1. Top left: Decisions of both participants over 300 iterations. Bottom left: Number $N_1(t)$ of 1-decisions over time t. Note that $N_1 = 1$ corresponds to the system optimum, while $N_1 = 2$ corresponds to the user equilibrium of the one-shot game. Right: Cumulative payoff of both players in the course of time t (i.e. as a function of the number of iterations). Once the coherent, oscillatory cooperation is established ($t > 220$), both individuals have high payoff gains on average.

Let us first focus on the two-person route-choice game (see Table 1). For our choice of parameters, the best individual payoff in each iteration is obtained by choosing route 1 (the "freeway") and have the co-player(s) choose route 2. Choosing route 1 is the dominant strategy of the one-shot

[3] I.e. our two-person experiment. In section 3 we give the corresponding game theoretical classification.

Hans-Ulrich Stark, Dirk Helbing, Martin Schönhof, and Janusz A. Hołyst:
Alternating cooperation strategies in a Route Choice Game:
Theory, experiments, and effects of a learning scenario.
In: A. Innocenti and P. Sbriglia (eds.): *Games, Rationality and Behaviour.*
Houndmills and New York: Palgrave McMillan, pp. 256–273, 2008.

game, and players are tempted to use it. This produces an initial tendency towards the "strongly stable" user equilibrium [40] with 0 points for everyone. However, this decision behavior is not Pareto efficient in the repeated game. Therefore, after many iterations, the players often learn to establish the Pareto optimum of the multi-stage supergame by selecting route 1 in turns (see Fig. 2). As a consequence, the experimental payoff distribution shows a maximum close to 0 points in the beginning and a peak at 100 points after many iterations (see Fig. 3), which clearly confirms that the choice behavior of test persons tends to change over time.

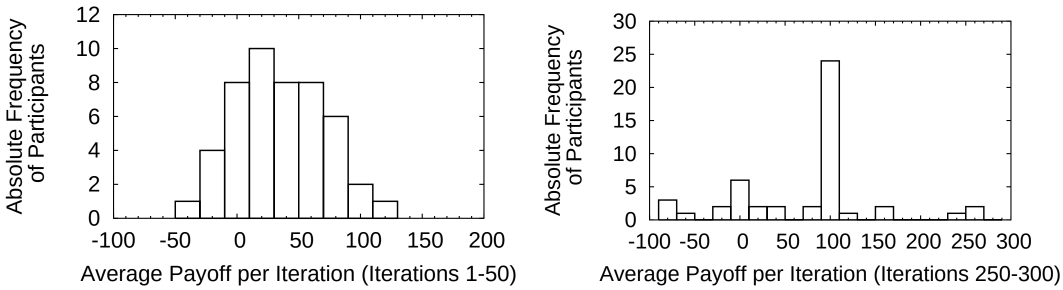

Figure 3: Frequency distributions of the average payoffs of the 48 players participating in our 24 two-person route choice experiments. Left: Distribution during the first 50 iterations. Right: Distribution between iterations 250 and 300. The initial distribution with a maximum close to 0 points (left) indicates a tendency towards the user equilibrium corresponding to the dominant strategy of the one-shot game. However, after many iterations, many individuals learn to establish the system optimum with a payoff of 100 points (right).

Nevertheless, in 7 out of 24 two-person experiments, persistent cooperation did not emerge during the experiment. Fig. 4 shows an example for that. In the next section 2.3, we will identify reasons for this.

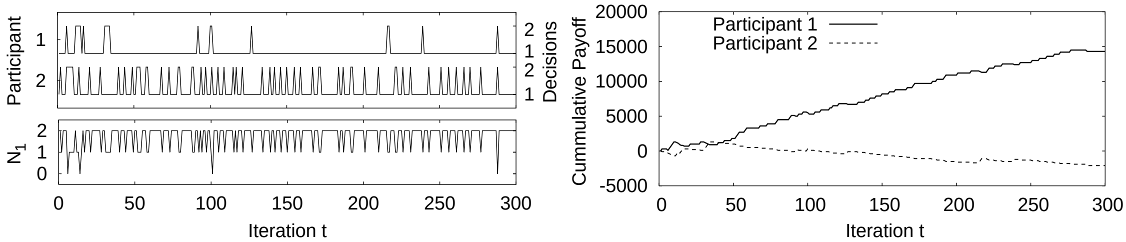

Figure 4: Representative example for a 2-person route choice experiment, in which no alternating cooperation was established. Due to the small changing frequency of participant 1, there were not enough cooperative episodes that could have initiated coherent oscillations. Top left: Decisions of both participants over 300 iterations. Bottom left: Number $N_1(t)$ of 1-decisions over time t. Right: The cumulative payoff of both players in the course of time t shows that the individual with the smaller changing frequency has higher profits.

2.3 Transition to cooperation

In this section, we will provide some crucial insights into the transition process to coherent, coordinated behavior. More details regarding this process and, in particular, a novel model of reinforcement learning that reproduces our experimental observations surprisingly well can be found in Helbing *et al.* [18].

2.3.1 Analytical estimation of the first cooperative episode

In the following, we will analytically estimate the time period until the first system optimal solution is explored. Therefore, we focus on the time period before persistent oscillatory cooperation is established and denote the occurrence probability that individual i chooses alternative $k \in \{1, 2\}$ by $P_i(k)$. The quantity $p_i(l|k)$ shall represent the conditional probability of choosing l in the next iteration, if k was chosen by person i in the present one. Assuming stationarity for reasons of simplicity, we expect the relationship

$$p_i(2|1)P_i(1) = p_i(1|2)P_i(2)\,, \tag{4}$$

i.e. the (unconditional) occurence probability $P_i(1, 2) = p_i(2|1)P_i(1)$ of having alternative 1 in one iteration and 2 in the next agrees with the joint occurence probability $P_i(2, 1) = p_i(1|2)P_i(2)$ of finding the opposite sequence 21 of decisions:

$$P_i(1, 2) = P_i(2, 1)\,. \tag{5}$$

Moreover, if r_i denotes the average changing frequency of person i until persistent cooperation is established, we have the relation

$$r_i = P_i(1, 2) + P_i(2, 1)\,. \tag{6}$$

Therefore, the probability that all N players simultaneously change their decision from one iteration to the next is $\prod_{i=1}^{N} r_i$. Note that there are 2^N such realizations of N decision changes 12 or 21, which all have the same occurence probability because of Eqn. (5). Among these, only the ones in which $N/2$ players change from 1 to 2 and the other $N/2$ participants change from 2 to 1 establish cooperative episodes, given that the system optimum corresponds to an equal distribution over both alternatives. Considering that the number of different possibilities of selecting $N/2$ out of N persons is given by the binomial coefficient, the occurence probability of a cooperative episode is

$$P_{\mathrm{c}} = \frac{1}{2^N} \left(\begin{array}{c} N \\ N/2 \end{array} \right) \prod_{i=1}^{N} r_i \tag{7}$$

Hans-Ulrich Stark, Dirk Helbing, Martin Schönhof, and Janusz A. Hołyst:
Alternating cooperation strategies in a Route Choice Game:
Theory, experiments, and effects of a learning scenario.
In: A. Innocenti and P. Sbriglia (eds.): *Games, Rationality and Behaviour.*
Houndmills and New York: Palgrave McMillan, pp. 256–273, 2008.

(at least in the ensemble average). Since the expected time period T until the cooperative state incidentally occurs equals the inverse of P_c, we finally find the formula

$$T = \frac{1}{P_c} = 2^N \frac{(N/2)!^2}{N!} \prod_{i=1}^{N} \frac{1}{r_i} . \tag{8}$$

This formula is well confirmed by our 2-person experiments (see Fig. 5). It gives the lower bound for the expected value of the minimum number of required iterations until persistent cooperation can spontaneously emerge (if already the first cooperative episode is continued forever).

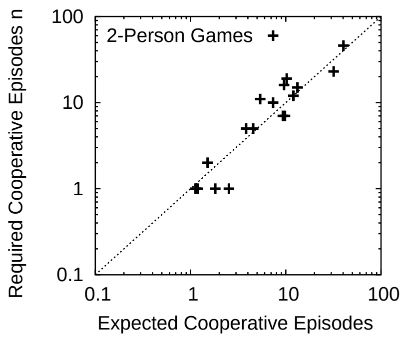

Figure 5: Comparison of the required number of cooperative episodes y (counted in the experiments as cooperative episodes that did not lead to persistent cooperation) with the expected number x of cooperative episodes (approximated as occurrence time of persistent cooperation, divided by the expected time interval T until a cooperative episode occurs by chance). Note that the data points support the relationship $y = x$ and, thereby, formula (8).

Obviously, the occurence of oscillatory cooperation is expected to take much longer for a large number N of participants. This tendency is confirmed by our 4-person experiments compared to our 2-person experiments. It is also in agreement with intuition, as coordination of more people is more difficult. (Note that mean first passage or transition times in statistical physics tend to grow exponentially in the number N of particles as well.)

Besides the number N of participants, another critical factor for the cooperation probability is the changing frequencies r_i. They are needed for the exploration of innovative strategies, coordination and cooperation. Although the instruction of test persons would have allowed them to conclude that taking turns would be a good strategy, the changing frequencies r_i of some individuals was so small that cooperation within the duration of the respective experiment did not occur, in accordance with formula (8). The unwillingness of some individuals to vary their decisions is sometimes called "conservative" [7, 44, 45] or "inertial behavior" [8]. Note that, if a player never reciprocates "offers" by other players, this may discourage further "offers" and reduce the changing frequency of the other player(s) as well (see the decisions 50 through 150 of player 2 in Fig. 2).

Our experimental time series show that most individuals initially did not know that a periodic decision behavior would allow them to establish the system optimum. This indicates that the required depth of strategic reasoning [12] and the related complexity of the game for an average person are already quite high, so that intelligence may matter. Compared to control experiments, the hint that the maximum average payoff of 100 points per round could be reached "by variable, situation-dependent decisions", increased the average changing frequency (by 75 percent) and with this the occurrence frequency of cooperative episodes. Thereby, it also increased the chance that persistent cooperation established during the duration of the experiment.

2.3.2 Strategy coefficients

In order to characterize the strategic behavior of individuals and predict their chances of cooperation, we have introduced some strategy coefficients. For this, let us introduce the following quantities, which are determined from the iterations before persistent cooperation is established:

- c_i^k = relative frequency of a *changed subsequent* decision of individual i, if the payoff was negative ($k = -$), zero ($k = 0$), or positive ($k = +$).
- s_i^k = relative frequency of individual i to *stay* with their previous decision, if the payoff was negative ($k = -$), zero ($k = 0$), or positive ($k = +$).

The Yule-coefficient

$$Q_i = \frac{c_i^- s_i^+ - c_i^+ s_i^-}{c_i^- s_i^+ + c_i^+ s_i^-} \tag{9}$$

with $-1 \leq Q_i \leq 1$ was used by Selten *et al.* [44] to quantify short term strategies of the participants. High absolute values of Q_i indicate a strong tendency to base decisions on the last received payoff, whereas values around zero negate such a correlation. Positive values $Q_i > 0.5$ correspond to a "win-stay lose-shift" strategy [34] and are referred to as "direct response" strategies. On the opposite, "contrarian response" is indicated by values $Q_i < -0.5$ and could be explained by the participants' conjecture that high payoffs will attract too many other players and low payoffs cause them to leave the chosen alternative.

However, a problem arises if one of the variables c_i^-, s_i^+, c_i^+, or s_i^- assumes the value 0. Then, we have $Q_i \in \{-1, 1\}$, independently of the other three values. If two of the variables become zero, Q_i is sometimes even undefined. Moreover, if the values are small, the resulting conclusion is not reliable. Therefore, we instead prefer to use the percentage difference

$$S_i = \frac{c_i^-}{c_i^- + s_i^i} - \frac{c_i^+}{c_i^+ + s_i^+}. \qquad (10)$$

Additionally, we introduce the proportion of changed decisions after the user equilibrium

$$Z_i = \frac{c_i^0}{c_i^0 + s_i^0} \qquad (11)$$

as a second strategy coefficient. The above described interpretation of Q_i also holds for our S_i, but comparisons with Q_i show that this time direct response corresponds to values $S_i > 0.25$, while values $S_i < -0.25$ correspond to contrarian behavior. The values of Z_i range from 0 to 1. A value $Z_l = 0$ would mean that individual i does not change routes, if the user equilibrium was reached. $Z_i = 1$ would would imply that person i always changes in the equilibrium case, while $Z_i \approx 0.5$ would correspond to a random response.

The significance of those strategic behaviors and their effect on the group performance in our experiments can be seen in Fig. 6.

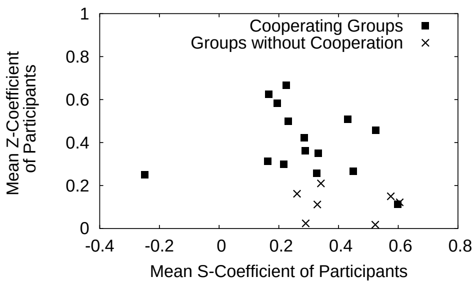

Figure 6: S- and Z-coefficients averaged over both participants in all 24 two-person route choice games. Final stages of persistent cooperation were excluded. The mainly small, but positive values of S indicate a slight tendency towards direct responses. However, the S-coefficient is barely significant for the emergence of persistent oscillations. A good indicator for their establishment is a sufficiently large Z-value.

2.4 Effects of the "learning scenario"

So far, especially in our two-person experiments, we have investigated the learning of participants over time (e.g. 300 iterations). It would be interesting to study the effect of past experience on the coordination process as it has been studied for the Prisoner's Dilemma by many publications!

For example, Andreoni and Samuelson [1] conducted a two-period Prisoner's dilemma experiment where the absolute payoff values can vary between the two periods[4]. One empirical result was that well adjusted lower amounts in the first period (and, therefore, higher amounts in the second period) lead to the highest overall payoffs in the experiments. That means that a two-period setup can induce a learning effect that increases the overall performance of the system.

Our approach described in section 2.1 is different, but not unrelated. According to intuition and formula (8), it is more difficult (and takes more iterations) to coordinate between many players. This is supported by experimental results indicating that the probability of persistent cooperation within a given number of iterations is decreasing with the Number N of participants (see Table 2). Although the optimal, time-dependent behavior is not more complicated in games with $N > 2$, the accidental occurrence of cooperative episodes is less likely.

Therefore, our hypothesis is that "positively experienced" players, who already have discovered and experienced successful cooperation in a transparent[5] setup of a two-person Route Choice Game, will cooperate in the following four-person experiments with the same payoff structure more easily.

treatment	groups/runs	coop. runs	payoff
2 unexp.	24	17	63.96
4 exp.	12	7	34.19
4 exp. (0)	2	0	-35.20
4 exp. (2)	3	0	-18.56
4 exp. (4)	7	7	76.63

Table 2: Overview of different treatments. We conducted experiments with 2 parallel runs of 2 unexperienced participants first ('2 unexp.') who played a 4-person game afterwards ('4 exp.'). The results of the 4-person games can be distinguished according to the number of participants who discovered persistent cooperation in the previous 2-person games (number shown in brackets). The third column shows the number of runs where persistent cooperation occurred and the last one contains the average payoffs per participant and per iteration.

Table 2 provides an overview of our results for different treatments. Although the standard deviations are very high, the mean values are still indicative of systematic effects. In spite of the limits imposed by only 12 data points coming from the 12 four-person experiments, the results are clear-cut:

- All of the 7 groups with only "positively experienced" participants reached persistent cooperation also in the four-person experiments.

[4] However, the proportions of the values within one payoff matrix and the sum of two subsequent payoff matrices were always constant.

[5] In the two-person experiments, the participants have sufficient information to recognize the actions of all players in the game. This does not hold for the four-person experiments.

Hans-Ulrich Stark, Dirk Helbing, Martin Schönhof, and Janusz A. Hołyst:
Alternating cooperation strategies in a Route Choice Game:
Theory, experiments, and effects of a learning scenario.
In: A. Innocenti and P. Sbriglia (eds.): *Games, Rationality and Behaviour.*
Houndmills and New York: Palgrave McMillan, pp. 256–273, 2008.

- None of the 5 remaining groups (in which 2 or more participants did not experience persistent cooperation before) reached persistent cooperation in the four-person experiments.
- In the 3 groups with 2 "positively experienced" participants only, the average group performance was higher than in the 2 groups without any "positively experienced" participants[6].

The first item indicates that our hypothesis is empirically evident. Although one could expect this result since the optimal, time-dependent strategic behavior is the same in both variants of the experiment, it is worth noticing that it would be profitable to be the only defector, and there is anonymity among the four participants. Furthermore, the second and the third item suggest that cooperative individuals are not able to transfer their knowledge of the cooperative strategy to players with no cooperation experience.

3 Game theoretical implications

	1	2
Route 1	0	300
Route 2	−100	−200

	11	12	21	22
11	0	300	300	600
12	−100	−200	200	100
21	−100	200	−200	100
22	−200	−300	−300	−400

Figure 7: Illustration of the concept of higher-order games defined by n-stage strategies. Left: Payoff matrix $\mathbf{P} = (P_{ij})$ of the one-shot 2x2 Route Choice Game. Right: Payoff matrix $\left(P^{(2)}_{(i_1 i_2),(j_1 j_2)}\right) = (P_{i_1 j_1} + P_{i_2 j_2})$ of the 2nd-order Route Choice Game defined by 2-stage decisions (right). The arrows indicate temptation (1), punishment (2), an offer (3), and reward for this offer (4). Note that the time-averaged payoff of this cycle lies below the system optimum.

As mentioned before, this situation of day-to-day route choice is also interesting from the game theoretical point of view. The left hand side of Fig. 7 gives the usual payoff matrix representation of our symmetrical *Route Choice Game*[7]. It is characterized by a dominant strategy which leads

[6]Note that these values are based on only 2 and 3 runs, respectively, and thus their difference may not be significant.

[7]The Route Choice Game is one of three *congestion games* among the symmetrical 2x2 games. Any congestion game is characterized by strictly non-increasing payoffs for every alternative with higher frequency of selection. Therefore, *Leader* and *Battle of the Sexes* (see Fig. 8) are the remaining two. Furthermore, congestion games belong to the class of *potential games* [31], for which many theorems are available.

Hans-Ulrich Stark, Dirk Helbing, Martin Schönhof, and Janusz A. Hołyst:
Alternating cooperation strategies in a Route Choice Game:
Theory, experiments, and effects of a learning scenario.
In: A. Innocenti and P. Sbriglia (eds.): *Games, Rationality and Behaviour.*
Houndmills and New York: Palgrave McMillan, pp. 256–273, 2008.

to a Pareto efficient Nash equilibrium in the one-shot game. This fulfilled efficiency criteria is a major difference to the Prisoner's Dilemma. However, since the sum of all payoffs in the off-diagonal (one player plays route 1, the other one chooses route 2) is highest, there is a strategical conflict in the repeated game.

Before we have a closer look on this conflict, let us classify the Route Choice Game in line with the other symmetrical 2x2 games. Since there is this symmetry between the payoffs of both players, we can define the game by giving 4 payoff values (cf. left of Fig. 7). Now, we make one more abstraction. Instead of cardinal values, we just take the ordinal ranking of the payoff values into account. Hence, we can fix the two outcomes of any game where both players take the same decision. For example, we fix the lower outcome of both at -200 and the higher at 0 (see Fig 8). As a consequence, the whole information regarding the ordinal ranking of the four payoff values are now given by the two remaining outcomes and their classification into smaller than -200, between -200 and 0, or higher than 0. On the right of Fig. 8 we extended the resulting concise classification of the games as it was used before by Eriksson and Lindgren [13].

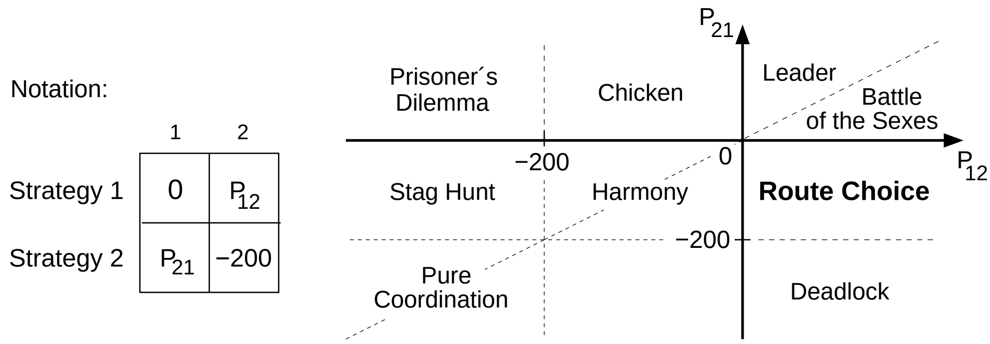

Figure 8: Classification of symmetrical 2x2 games according to their payoffs P_{ij}. Two payoff values have been kept constant as payoffs may be linearly transformed and the two strategies of the one-shot game renumbered. Our choice of $P_{11} = 0$ and $P_{22} = -200$ was made to define a payoff of 0 points in the user equilibrium and an average payoff of 100 in the system optimum of our investigated route choice game with $P_{12} = 300$ and $P_{21} = -100$.

According to the game-theoretical literature, there are 12 ordinally distinct, symmetric 2x2 games [40], but after excluding strategically trivial games in the sense of having equilibrium points that are uniquely Pareto-efficient, there remain four archetypical 2x2 games: the Prisoner's Dilemma, the Battle of the Sexes, Chicken (Hawk-Dove), and Leader [39]. However, this conclusion is only correct, if the four payoff values P_{ij} are specified by the four values $\{1, 2, 3, 4\}$. Taking different values would lead to a different conclusion: If we name subscripts so that $P_{11} > P_{22}$, a strategical conflict between a user equilibrium and the system optimum results when

$$P_{12} + P_{21} > 2P_{11} \,. \tag{12}$$

Our conjecture is that players tend to develop alternating forms of reciprocity if this condition is fulfilled, while symmetric reciprocity is found otherwise. This has the following implications (see Fig. 8):

- If the 2x2 games Stag Hunt, Harmony, or Pure Coordination are repeated frequently enough, we always expect a symmetrical form of cooperation.
- For Leader and the Battle of the Sexes, we expect the establishment of asymmetric reciprocity, as has been found by Browning and Colman with a computer simulation based on a genetic algorithm incorporating mutation and crossing-over [9].
- For the games Route Choice, Deadlock, Chicken, and Prisoner's Dilemma both, symmetric (simultaneous) and asymmetric (alternating) forms of cooperation are possible, depending on whether condition (12) is fulfilled or not. Note that this condition cannot be met for some games, if one restricts to ordinal payoff values $P_{ij} \in \{1, 2, 3, 4\}$ only. Therefore, this interesting problem has been largely neglected in the past (with a few exceptions, e.g. [32, 35]). In particular, convincing experimental evidence of alternating reciprocity is missing so far.

Finally, let us have a look at the advantages of the concept of *higher order games*, illustrated on the right of Fig. 8. There, we can clearly see the strategical conflict our test persons faced due to the many iterations played. Still, strategy 11 is strict dominant and, therefore, always taking route 1 is still the only Nash solution. But it is not any longer Pareto efficient as an encounter of strategy 12 with 21 yields 200 payoff points for both players. This second order Route Choice Game can exactly be divided into two subconflicts.

First, assume strict readiness to cooperate for both players. That means that the temptation arrow (1) in Fig. 7 has no meaning in that case. Then, only strategies 12 and 21 would come into consideration to the players. The remaining game, consisting of the four possible solutions in the inset of the 4x4 matrix, is an "anti-coordination" problem, but allows for a symmetrical form of cooperation.

Second, let us neglect this problem of anti-coordination. Say, if both players choose one of these cooperative strategies, an "invisible coordinator" guarantees the optimal outcome. What remains is the typical Prisoner's Dilemma situation with two identical cooperation strategies and one additional option that can be neglected since it is strictly dominated.

4 Conclusions and outlook

In this paper, the main results of a former work by the authors are reviewed [18]. We abstract real traffic phenomena into an experimental setup in order to observe self-organized coordination in favour of the non-trivial system optimal solution. The latter is characterized by a very unfair distribution of payoffs among the participants in our computer experiments. Therefore, the requirements for successful, long-term cooperation are quite high and can be described by three points: (i) exploration of coherent, alternating decisions, either intentionally or by chance, (ii) readiness to cooperate instead of short-term maximization, (iii) ability to coordinate with the co-player(s), as uncoordinated alternations even worsen both the system's and the individual performance.

The difference with respect to the previous work on route choice experiments is as follows: Selten *et al.* [44] have studied a setup in which the system optimum almost agreed with the user equilibrium. Therefore, their experimental results do not allow to distinguish between them. The same applies to Helbing *et al.* [17]. They have varied the payoff parameters in time and tested different information treatments, including scenarios with route choice recommendations. These experiments were performed with $N \geq 9$ test persons. In contrast, this paper focusses on the case of 2- and 4-person games with constant payoff parameters and a pronounced difference between the system optimum and the user equilibrium.

In our data analysis, we do not only document the emergence of coherent, alternating cooperation strategies, but also investigate details of the transition process from initial tendencies towards the inefficient user equilibrium to phase-coordinated alternations. Especially the individual Z-coefficients, that quantify reactions of the participants on the equilibrium outcome, are shown to be a crucial factor for or against cooperation. Furthermore, the presented results regarding the effects of our learning scenario suggest that the coordination process towards persistent cooperation may be the bigger problem compared to the evolution of cooperativeness. Once the innovation of asymmetric cooperation happened and its long-term profitability is experienced, there is a very good chance of continuing this strategical behavior in more complex environments. This fact leads us to the conclusion that investigating the Prisoner's Dilemma may be not enough for resolving social dilemmas as they sometimes do not provide symmetrical solutions.

Therefore, the emphasis of this paper lays on game theoretical relevance of the Route Choice Game and the not very common study of asymmetric, alternating cooperation. Especially the latter calls for further investigations.

Appendix

Instructions for participants

In this Route Choice experiment, every participant has to repeatedly choose between two possible routes. All have the same origin and destination.

- Route A corresponds to a freeway (may be fast or congested)
- Route B corresponds to another road.

After each round you will be informed about your payoff for this round. The payoff only depends on your last chosen route and the number of participants who have chosen this route as well. The less participants, the more payoff points you get (i.e. the faster it was possible to drive on this road).

If 2 [3] out of 2 [4] participants choose route A, everyone gets 0 points.
If 1 [2] out of 2 [4] participants choose route A, the average payoff for all participants is 100 points, but A-choosers profit at the cost of B-choosers.

Nevertheless, every participant can reach up to 100 points per decision on average. The sum of your payoff points after 300 rounds will determine your payment at the end of the experiment (100 payoff points = 0.01 EUR).

Acknowledgements

H.-U. S. thanks João F. M. Rodrigues for numerous scientific discussions that will surely lead to further interesting results soon.

References

[1] Andreoni, J. and Samuelson, L., Building rational cooperation, *Journal of Economic Theory* **127**, 117–154 (2006).

[2] Arthur, W. B., Inductive reasoning and bounded rationality, *American Econonmic Review* **84**, 406–411 (1994).

[3] Axelrod, R. and Dion, D., The further evolution of cooperation, *Science* **242**, 1385–1390 (1988).

[4] Axelrod, R. and Hamilton, W. D., The evolution of cooperation. *Science* **211**, 1390–1396 (1981).

[5] Binmore, K. G., Evolutionary stability in repeated games played by finite automata, *Journal of Economic Theory* **57**, 278–305 (1992).

[6] Binmore, K., *Fun and Games: A Text on Game Theory* (Heath, Lexington, MA, 1992), pp. 373–377.

[7] Bohnsack, U., Uni DuE: Studie SURVIVE gibt Einblicke in das Wesen des Autofahrers, Press release by *Informationsdienst Wissenschaft* (January 21, 2005).

[8] Bottazzi, G. and Devetag, G., Coordination and self-organization in minority games: Experimental evidence, Working Paper 2002/09, Sant'Anna School of Advances Studies, May 2002.

[9] Browning, L. and Colman, A. M., Evolution of coordinated alternating reciprocity in repeated dyadic games, *J. Theor. Biol.* **229**, 549–557 (2004).

[10] Challet, D. and Marsili, M., Relevance of memory in minority games, *Phys. Rev. E* **62**, 1862–1868 (2000).

[11] Challet, D. and Zhang, Y.-C., Emergence of cooperation and organization in an evolutionary game, *Physica A* **246**, 407–418 (1997).

[12] Colman, A. M., Depth of strategic reasoning in games, *TRENDS in Cognitive Sciences* **7**(1), 2–4 (2003).

[13] Eriksson, A. and Lindgren, K. Cooperation in an unpredictable environment, in *Proceedings of Artificial Life VIII* (eds. Standish, R. K., Bedau, M. A. and Abbass, H. A.) 394–399 (MIT Press, Sidney, 2002) and poster available at http://frt.fy.chalmers.se/cs/people/eriksson.html

[14] Glance, N. S. and Huberman, B. A., The outbreak of cooperation, *Journal of Mathematical Sociology* **17**(4), 281–302 (1993).

[15] Greenshield, B. D., A study of traffic capacity, in *Proceedings of the Highway Research Board* (Highway Research Board, Washington, D.C.), Vol. 14 (1935), pp. 448–477.

[16] Hardin, G., The tragedy of the commons, *Science* **162**, 1243–1248 (1968).

[17] Helbing, D., Schönhof, M., and Kern, D., Volatile decision dynamics: Experiments, stochastic description, intermittency control, and traffic optimization. *New Journal of Physics* **4**, 33.1–33.16 (2002).

[18] Helbing, D., Schönhof, M., Stark, H.-U., and Hołyst, J.A., How individuals learn to take turns: Emergence of alternating cooperation in a congestion game and the Prisoner's Dilemma. *Advances in Complex Systems* **8**(1), 87–116 (2005).

[19] Hofbauer, J. and Sigmund, K., *The Theory of Evolution and Dynamical Systems* (Cambridge University Press, Cambridge, 1988).

[20] Hogg, T. and Huberman, B. A., Controlling chaos in distributed systems, *IEEE Transactions on System, Man, and Cybernetics* **21**(6), 1325–1333 (1991).

[21] Iwanaga, S. and Namatame, A., The complexity of collective decision, *Nonlinear Dynamics, Psychology, and Life Sciences* **6**(2), 137–158 (2002).

[22] Kephart, J. O., Hogg, T., and Huberman, B. A., Dynamics of computational ecosystems, *Phys. Rev. A* **40**(1), 404–421 (1989).

[23] Lo, T. S., Chan, H, Y., Hui, P. M., and Johnson, N. F., Theory of networked minority games based on strategy pattern dynamics, *Phys. Rev. E* **70**, 056102 (2004).

[24] Lo, T. S., Hui, P. M., and Johnson, N. F., Theory of the evolutionary minority game, *Phys. Rev. E* **62**, 4393–4396 (2000).

[25] Macy, M. W. and Flache, A., Learning dynamics in social dilemmas, *Proceedings of the National Academy of Sciences USA* **99**, Suppl. 3, 7229–7236 (2002).

[26] Mansilla, R., Algorithmic complexity in the minority game, *Phys. Rev. E* **62**, 4553–4557 (2000).

[27] Marsili, M., Mulet, R., Ricci-Tersenghi, F., and Zecchina, R., Learning to coordinate in a complex and nonstationary world, *Phys. Rev. Lett.* **87**, 208701 (2001).

[28] McNamara, J. M., Barta, Z., and Houston, A. I., Variation in behaviour promotes cooperation in the prisoner's dilemma game, *Nature* **428**, 745–748 (2004).

[29] Michor, F. and Nowak, M. A., The good, the bad and the lonely, *Nature* **419**, 677–679 (2002).

[30] Milinski, M., Semmann, D., and Krambeck, H.-J., Reputation helps solve the 'tragedy of the commons', *Nature* **415**, 424–426 (2002).

[31] Monderer, D. and Shapley, L. S., Potential games, *Games and Economic Behavior* **14**, 124–143 (1996).

[32] Neill, D. B., Cooperation and coordination in the turn-taking dilemma, *In TARK*, 231–244 (2003).

[33] Nowak, M. A., Sasaki, A., Taylor, C., and Fudenberg, D., Emergence of cooperation and evolutionary stability in finite populations, *Nature* **428**, 646–650 (2004).

[34] Nowak, M. A. and Sigmund, K., A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game, *Nature* **364**, 56–58 (1993).

[35] Nowak, M. A. and Sigmund, K. The alternating prisoner's dilemma, *J. theor. Biol.* **168**, 219–226 (1994).

[36] Nowak, M. A. and Sigmund, K., Evolution of indirect reciprocity by image scoring, *Nature* **393**, 573–577 (1998).

[37] Pigou, A. C., *The Economics of Welfare* (Macmillan, London, 1920).

[38] Queller, D. C., Kinship is relative, *Nature* **430**, 975–976 (2004).

[39] Rapoport, A., Exploiter, leader, hero, and martyr: the four archtypes of the 2x2 game, *Behav. Sci.* **12**, 81–84 (1967).

[40] Rapoport, A. and Guyer, M., A taxonomy of 2x2 games, *Gen. Systems* **11**, 203–214 (1966).

[41] Riolo, R. L., Cohen, M. D., and Axelrod, R., Evolution of cooperation without reciprocity, *Nature* **414**, 441–443 (2001).

[42] Schelling, T. C., *Micromotives and Macrobehavior* (WW Norton and Co, New York, 1978), pp. 224–231+237.

[43] Schweitzer, F., Behera, L., and Mühlenbein, H., Evolution of cooperation in a spatial prisoner's dilemma, *Advances in Complex Systems* **5**(2/3), 269–299 (2002).

[44] Selten, R. *et al.*, Experimental investigation of day-to-day route-choice behaviour and network simulations of autobahn traffic in North Rhine-Westphalia, in *Human Behaviour and Traffic Networks*, eds. Schreckenberg, M. and Selten, R. (Springer, Berlin, 2004), pp. 1–21.

[45] Selten, R., Schreckenberg, M., Pitz, T., Chmura, T., and Kube, S., Experiments and simulations on day-to-day route choice-behaviour, see http://papers.ssrn.com/sol3/papers.cfm?abstract_id=393841.

[46] Semmann, D., Krambeck, H.-J., and Milinski, M., Volunteering leads to rock-paper-scissors dynamics in a public goods game, *Nature* **425**, 390–393 (2003).

[47] Szabó, G. and Hauert, C., Phase transitions and volunteering in spatial public goods games, *Phys. Rev. Lett.* **89**, 118101 (2002).

[48] Wardrop, J. G., Some theoretical aspects of road traffic research, in *Proceedings of the Institution of Civil Engineers II*, Vol. 1 (1952), pp. 325–378.

[49] Yuan, B. and Chen, K., Evolutionary dynamics and the phase structure of the minority game, *Phys. Rev. E* **69**, 067106 (2004).