



## UvA-DARE (Digital Academic Repository)

### Evolution of altruism: Exploring adaptive landscapes

Spichtig, M.

**Publication date**  
2013

[Link to publication](#)

#### **Citation for published version (APA):**

Spichtig, M. (2013). *Evolution of altruism: Exploring adaptive landscapes*. [Thesis, fully internal, Universiteit van Amsterdam].

#### **General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### **Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

## When and how does mutation-generated variation promote the evolution of cooperation?

Mathias Spichtig · Maurice W. Sabelis · Martijn Egas

**Abstract** It has been hypothesized that mutation-generated variation in behavior can drive the evolution of cooperation. In this article, we distinguish between two effects of mutation that can generate this in evolutionary games of the finitely repeated Prisoner's Dilemma in infinite asexual populations. First, we consider a particular setting with strategies that play Tit-for-Tat for the first  $x$  rounds and unconditionally defect thereafter, and where mutation only alters  $x$ . For this regime, we show how cooperation can evolve through the direct effect of mutation, i.e., the instantaneous impact of mutational variation before selection may act upon these mutants. This direct effect suffices to explain earlier findings that behavioral variation can promote the evolution of cooperation. However, we question the generality of these findings because for many (if not most) mutation regimes mutation generates the highest direct effect on unconditional defectors (*AllD*). Overall, mutation is therefore expected to hamper rather than promote the evolution of cooperation. We identify special conditions (e.g., intermediate mutation rates) for which this negative direct effect of mutation on cooperation can be countered by an indirect effect of mutation, i.e., the fitness impact that individuals obtain from interactions with descendants of mutants. Simulations for these conditions show that populations dominated by *AllD* are invaded by cooperative strategies despite the positive direct effect of mutation on *AllD*. Thus, here it is the indirect effect of mutation on cooperative strategies that drives the evolution of cooperation. The population evolves towards a state where the action 'cooperate' is more frequently executed, but this is not achieved by individuals triggering reciprocity ('genuine altruism'), but by individuals exploiting the willingness of others to cooperate ('exploitative altruism'). By highlighting the distinction between direct and indirect effect of mutation we provide a new perspective on how mutation-generated variation alters frequency-dependent selection.

## 1 Introduction

Evolution is driven by selection acting on heritable phenotypic variation. The amount of phenotypic variation can be described as a function of selection modulating the variation generated by mutation, recombination, and environmental factors. If selection is frequency-dependent, it can in turn be described as a function of the distribution of phenotypic variants in the population. A standard approach in evolutionary ecology is to ignore selection effects of phenotypic variation by reducing the analysis to interaction dynamics of an invader phenotype in an otherwise homogeneous resident population (Abrams 2001). This approach is defensible if the addition of other phenotypic variants constitutes no more than noise on the selective interactions between these two phenotypes. Under a regime of frequency-dependent selection, however, increasing phenotypic variation may alter the outcome of evolutionary analyses qualitatively. For example, increasing (mutation-generated) variation in behavior can promote the evolution of cooperation (McNamara et al. 2004, 2008; Eriksson and Lindgren 2005; Traulsen et al. 2009; McNamara and Leimar 2010).

The existence of behavioral variation is generic in studies of the evolution of cooperation (e.g., Wedekind and Milinski 1996; Fischbacher et al. 2001; Camerer 2003; Camerer and Fehr 2006; Henrich et al. 2006; Herrmann et al. 2008). Hence, it may provide a general explanation for the evolution of cooperation. However, this prediction by McNamara et al. (2004, 2008), Eriksson and Lindgren (2005), and Traulsen et al. (2009) is generated from simulating asexual populations. Each of these simulations is based on a particular choice of mutation regime, i.e., a mutation rate ( $\mu$ ), a set of strategies that mutation can generate, and a rule of how mutant strategies tend to differ from the parental strategies (variation from recombination and from environmental factors is ignored). Given the relative lack of knowledge on the genetic underpinning of behavioral traits, arbitrary choices are inevitable. In asexual organisms, mutation is the source of heritable variation and the mutation regime determines the expressed (heritable) variation. Therefore, we investigate whether variation-promoted evolution of cooperation is robust to the choice of mutation regime.

To analyze the effects of mutation regime on the evolutionary dynamics we start from a narrow-sense definition of mutants: mutants are the individuals with a genotype distinct from their parent, i.e., a mutation event occurred (transforming its genotype) during the generation of the individual. As a consequence of this narrow-sense definition, faithfully replicated offspring of mutants are not mutants themselves (in contrast to the use of the term in ‘mutant-wild type’ contexts where faithfully replicated offspring of mutants remain categorized as the mutant type). Because selection acts on variation after it is generated by mutation, mutants constitute the fraction of the population that has not yet been affected by selection. We refer to the direct effect of mutation when it concerns the average fitness contributions ( $\pi$ ) generated from interactions with the narrow-sense mutants. Of course, the fitness impact of mutation goes beyond that of the mutants alone as long as they (faithfully) produce descendants. We refer to this impact as the indirect effect of the mutants when it concerns the average fitness contributions ( $\theta$ ) generated from interaction with non-mutants (each with an ancestry of variable length to a mutant). Thus,  $\pi$  refers to the direct effect of mutation, i.e., the fitness impact of the mutation-generated variation yet unaffected by selection, and  $\theta$

refers to the indirect effect of mutation, i.e., the fitness impact of the mutation-generated variation after being subject to selection.

In this paper we show that the distinction between direct and indirect fitness effects is useful in providing insight in the impact of mutation-generated variation on the evolution of cooperation. Our results show that (1) earlier findings on the impact of behavioral variation on the evolution of cooperation are no more than special cases where there is a positive direct effect of mutation, (2) cooperation can evolve through an indirect effect of mutation-generated variation even if unconditional defectors benefit most from the direct effect, and (3) in cases with positive indirect effects of mutation, the selected strategies in the population exhibit exploitation of cooperative acts.

## 2 Model

We use the evolutionary game version of the finitely repeated Prisoner's Dilemma game (frPD; Nachbar 1992; Cressman 1996) to study the impact of different mutation regimes. In the Prisoner's Dilemma (PD) game, individuals engage in pairwise interactions where they choose to cooperate (**C**) or to defect (**D**). Both players execute their action simultaneously. Individuals receive payoff dependent on their own choice as well as the choice of their opponent, as given in the following payoff matrix:

|          |          |          |
|----------|----------|----------|
|          | <b>C</b> | <b>D</b> |
| <b>C</b> | <b>R</b> | <b>S</b> |
| <b>D</b> | <b>T</b> | <b>P</b> |

The payoffs follow the relations  $T > R > P > S$  and  $2R > T + S$ . The consequence is that defection generates higher individual payoffs  $T (> R)$  and  $P (> S)$ , while mutual cooperation maximizes the common payoff. In the frPD game, the Prisoner's Dilemma game is repeated for a fixed number ( $= r$ ) of times. The total number of deterministic strategies adjusted to an frPD game is finite (i.e.,  $2^{2r-1}$ ; Cressman 1996). Cressman (1996) provides a general analysis of evolutionary frPD games in absence of mutation. The strategy that defects ( $= d$ ) and the strategy that cooperates ( $= c$ ) are the two strategies in a one-shot game ( $r = 1$ ). It follows from the payoff matrix that for any population composition of the two strategies, individuals with strategy  $d$  always generate a higher average payoff than individuals with strategy  $c$ . Consequently, strategy  $d$  evolves towards fixation in infinite populations, i.e., the state in which mutual defection is observed in all PD games. For much the same reason, in mutation-free evolutionary frPD games ( $r > 1$ ) polymorphisms composed of all deterministic strategies evolve towards states in which the players exclusively defect (Nachbar 1992; Cressman 1996). Unconditional defectors (*AllD*) obtain fitness dominance ( $w_{AllD} \geq w_i$  for all strategies  $i$ ;  $w_i$  is the fitness of strategy  $i$ ) during this process of convergence towards full defection (Cressman 1996).

In our model, we assume infinite, asexual populations composed of deterministic strategies. The code used for the strategies is described below. Strategy  $i$  individuals generate average payoff  $\bar{p}_i = \sum_{j \in S} p_{ij} f_j$  from game interactions;  $p_{ij}$  is the payoff that strategy  $i$  individuals generate in interactions with strategy  $j$  individuals,  $f_j$  is the frequency of strategy  $j$ , and  $s$  represents the strategy set (i.e., the collection of the consid-

ered strategies). The average fitness of strategy  $i$  individuals is defined by  $w_i = \bar{p}_i + K$  where background fitness  $K$  is the game-unassociated fitness component (in the simulations, we used positive integers for  $K$ ) which is identical for each individual. Fitness determines reproductive success but not survival abilities. Consequently, all players have, independent of their performance, the same expected number of pairwise interactions over their life time in the game.

We analyze evolutionary changes in this model using both discrete generations and continuous (overlapping) generations. In the discrete-generation model (*dgm*), the frequency  $f_i'$  of strategy  $i$  in the next generation is determined as

$$f_i' = (1 - \mu) f_i \frac{w_i}{\bar{w}} + \mu u_i,$$

where  $\bar{w} = \sum_i f_i w_i$  is the average fitness of the population,  $\mu$  is the mutation rate, and  $u_i$  is the fraction of mutants that carry strategy  $i$ . The frequency dynamics in the continuous-generation model (*cgm*) follows the replicator-mutator dynamics (Page and Nowak 2002):

$$\dot{f}_i = f_i((1 - \mu)w_i - \bar{w}) + \mu u_i.$$

## 2.1 Analysis

We use the characteristic  $w_{AllD} \geq w_i$  of the evolutionary attractor to set a benchmark for the evolutionary effect of mutation: recurrent mutation significantly affects evolutionary frPD games whenever the fitness relation  $w_i > w_{AllD}$  is either persistently or periodically observed for at least one strategy  $i$ . Behavioral differences cause the differences in fitness between strategies  $i$  and *AllD*. Therefore,  $w_i > w_{AllD}$  implies – as *AllD* is the strategy which always defects – that strategy  $i$  employs at least some cooperation. Beyond that the observation  $w_i > w_{AllD}$  does not carry information about the frequency with which strategy  $i$  or the remainder of the population executes cooperation. However, frequencies of cooperative behaviors may positively correlate with mutation rates without challenging the fitness dominance of *AllD* (Willensdorfer and Nowak 2005). For example, mutation could frequently generate the strategy unconditional cooperator (*AllC*). The expected outcome of an increase in mutation rate would then be an increased execution of cooperation but also an increased fitness of *AllD*. If significant effects ( $w_i > w_{AllD}$ ) and increased cooperation co-occur, we take this as an indication that cooperation evolves due to a change in the direction of selection. Note that significant effects might emerge without the consequence that cooperation is amply executed.

## 2.2 Statistics

In our simulations, we sample the following behavioral statistics. The average cooperation in round  $x$  is given by  $\bar{C}_x = \sum_{i,j \in S} a_{x,ij} f_i f_j$  whereby  $a_{x,ij} = 1$  if strategy  $i$  cooperates

in round  $x$  against strategy  $j$  and  $a_{x,ij} = 0$  otherwise. The average cooperation per frPD game is given by  $\overline{C} = \sum_{i=1}^r \overline{C}_i$  (in case of non-equilibrium dynamics,  $\overline{C}$  is averaged over specified ranges of generations). As  $0 \leq \overline{C} \leq r$ , a population with  $\overline{C} \sim 1$  is interpreted as fairly cooperative if  $r = 1$  and as fairly uncooperative if  $r = 100$ . For comparisons of evolutionary frPD with different  $r$ -values, we thus use the average number of  $C$  executions per Prisoner's Dilemma game,  $\frac{\overline{C}}{r}$ . The average payoff per frPD game generated with payoff  $\mathbf{P}$  is given by  $\overline{p_P} = P \sum_{i,j \in S} a_{P,ij} f_i f_j$  whereby  $a_{P,ij}$  is the number of times strategy  $i$  generates payoff  $\mathbf{P}$  from strategy  $j$ . For payoffs  $\mathbf{T}$ ,  $\mathbf{S}$ , and  $\mathbf{R}$ , we analogously define the averages  $\overline{p_T}$ ,  $\overline{p_S}$ , and  $\overline{p_R}$ .

### 2.3 Strategies

We assume genotypes determine strategies. In certain mutation regimes neutral mutations can occur, i.e., parent and mutant offspring do not share the genotype but they do share the strategy. The mutation rate  $\mu$  determines the fraction of offspring that do not carry the parental genotype. We refer to these offspring as mutants. We define  $u_i$  as the fraction of mutants that express strategy  $i$ . The distributions of such fractions ( $u$ ) are constants if the probabilities with which mutants express any of the strategies are independent of the parental strategy. The  $u$  distributions are dependent variables of population compositions otherwise. We refer to the latter as variable  $u$  distributions.

Strategies are defined by their responses to each possible sequence of actions of an opponent in the game. We encode strategies by bit strings of digits  $\{d, c\}$  which represent actions  $\{\mathbf{D}, \mathbf{C}\}$  in response to the perceived actions of an opponent. The first position of the code gives the initial action. Second and third positions represent responses to the initial action  $\mathbf{D}$ , respectively  $\mathbf{C}$ , of an opponent. Fourth to seventh position represent responses to action sequences  $\{\mathbf{DD}, \mathbf{DC}, \mathbf{CD}, \mathbf{CC}\}$ , eighth to fifteenth positions represent responses to action sequences  $\{\mathbf{DDD}, \mathbf{DDD}, \mathbf{DCD}, \mathbf{DCD}, \mathbf{CDD}, \mathbf{CDC}, \mathbf{CCD}, \mathbf{CCC}\}$ , etc. As an example of the game with  $r = 3$ , the strategy  $cddcdcc$  cooperates in the first round, unconditionally defects in the second round, and defects in the third round only if the opponent defected in the first round and cooperated in the second round (i.e., played action sequence  $\mathbf{DC}$ ). In notations of strategy groups, we use dots to mark code positions at which the strategies of the group can differ, i.e., in this notation a dot can be replaced with either  $d$  or  $c$ . A special example is formed by the strategies  $(\{r1, r2, r3, r4\} \rightarrow \{d, dd., dd.d..., dd.d...d.....\})$  which signify the groups of strategies that exclusively defect in interactions among each other. We call these strategies 'defectors'. Populations exclusively composed of defectors are the only type of populations in which action  $\mathbf{C}$  is never executed (i.e., full defection by all players).

With the code explained above all deterministic frPD strategies can be described. We use the name  $rX$  for the strategy sets that contain all deterministic strategies, whereby  $X$  represents the number of rounds of the frPD game. Simulations comprise the  $rX$  sets  $\{r1, r2, r3, r4\}$  which contain respectively  $\{2, 8, 128, 32768\}$  strategies. For technical reasons due to the size of the set, only  $dgm$ -simulations were performed for the  $r4$  set.

Besides the  $rX$  sets, we use a type of strategy sets proposed by Nachbar (1992). We term these sets ‘ $TfTx$ ’. A strategy contained in ‘ $TfTx$ ’ sets is ‘Tit for Tat’ ( $TfT$ ; Axelrod and Hamilton 1981) which starts with  $C$  and thereafter repeats the previous actions of opponents. The strategy  $TfTx$  behaves as  $TfT$  in the first  $x$  rounds and unconditionally defects in the remaining  $(r - x)$  rounds. ‘ $TfTx$ ’ sets contain all  $TfTx$  strategies with  $x = \{0, 1, \dots, r - 1, r\}$ . The extremes represent  $AllD$  ( $x = 0$ ) and  $TfT$  ( $x = r$ ). In absence of mutation,  $AllD$  evolves towards fixation in any polymorphism of the ‘ $TfTx$ ’ strategies (Nachbar 1992). McNamara et al. (2004) altered the frPD game in that players can end the game after any round (whereby both players receive an identical payoff for each unplayed round). We do not implement this altered frPD game to avoid the consequent increase in the number of deterministic strategies and to take the work of Nachbar (1992) and Cressman (1996) as a starting point of our analysis (assuming it can be extrapolated to the game of McNamara et al. 2004). Hence, we consider ‘ $TfTx$ ’ sets as analogous to the sets used by McNamara et al. (2004). In order to confirm this analogy, we repeated the simulations of McNamara et al. (2004) using ‘ $TfTx$ ’ sets and retrieved qualitatively the same results in these simulations as McNamara et al. (2004) did with their sets (Spichtig, data not shown).

### 3 Results

In absence of recurrent mutation,  $AllD$  obtains fitness dominance at the evolutionary attractors for ‘ $TfTx$ ’ strategy sets (Nachbar 1992) and for  $rX$  strategy sets (Nachbar 1992; Cressman 1996). Hence, in our analyses we use the fitness of  $AllD$  as the benchmark for measuring the relative success of other strategies when mutation-generated variation is introduced in the model. Specifically, the evolutionary impacts of recurrent mutation are significant whenever strategies persistently or periodically exceed the fitness of  $AllD$  ( $w_i > w_{AllD}$ ). In the first subsection, we analyze the contributions to the significant impacts that result from direct effects of mutation and from the indirect effects of mutation-induced population compositions. In the second subsection, simulations of  $rX$ -populations are used to assess the relevance of indirect effects in absence of direct effects.

#### 3.1 Effects of mutation on the evolution of ‘ $TfTx$ ’ sets and of $rX$ sets

Populations can be subdivided in fractions  $\mu$  of mutants and  $(1 - \mu)$  of non-mutants. Strategy  $i$  generates average payoff  $\pi_i = \sum_{j \in s} u_j p_{ij}$  from interactions with mutants. The average fitness of strategy  $i$  individuals is  $w_i = (1 - \mu)\theta_i + \mu\pi_i + K$ , where  $\theta_i$  is the average payoff generated from interactions with non-mutants. We interpret the occurrence of inequalities  $\pi_i > \pi_{AllD}$  as direct effects and of inequalities  $\theta_i > \theta_{AllD}$  as indirect effects if they coincide with the observation of significant effects ( $w_i > w_{AllD}$ ). Direct effects and indirect effects are not mutually exclusive. We expect indirect effects generally to co-emerge with direct effects. In the following, the discussion on indirect effects focuses on their emergence for cases when direct effects are excluded (i.e.,  $\pi_{AllD} \geq \pi_i$  for all  $i \in s$  and at all population compositions).

For constant  $u$  distributions, strategies generate fixed returns  $\pi_i$ , i.e., the payoff from interactions with mutants is independent of the population composition. Then, direct effects can be excluded if  $(\pi_i - \pi_{AllD}) \leq 0$  for all strategies  $i$ . Direct effects can occur (and are inevitable for sufficiently high  $\mu$ -values) if  $(\pi_i - \pi_{AllD}) > 0$  for at least one strategy  $i$ . For variable  $u$  distributions, the averages  $\pi_i$  are functions of population compositions. These compositions are also functions of the mutation rate  $\mu$ . As a consequence, there is no simple expression for when direct effects can emerge. We focus on analytical results assuming constant  $u$  distributions and only briefly discuss the more complicated case of variable  $u$  distributions.

For ‘ $TfTx$ ’ sets, the difference in performance between  $TfTx$  and  $AllD$  in interactions with mutants can be given by the recursion  $\pi_{TfTx} - \pi_{AllD} = \sum_{i=0}^{x-1} (\pi_{TfTi+1} - \pi_{TfTi})$  (Note that in ‘ $TfTx$ ’ notation,  $AllD$  is  $TfT0$ ). The adjacent strategies  $TfTx-1$  and  $TfTx$  perform identically with mutants expressing strategies  $TfTy$  for which  $y \leq x-2$ .  $TfTx$  individuals generate one additional round of mutual cooperation from interactions with mutants expressing strategies  $TfTy$  for which  $y \geq x$ .  $TfTx$  individuals are exploited by  $TfTx-1$  mutants at a single occasion, and they do not exploit  $TfTx$  mutants in round  $x+1$ . Consequently, strategy  $TfTx$  is more effective in interactions with mutants than  $TfTx-1$  ( $\pi_{TfTx} - \pi_{TfTx-1} > 0$ ) if

$$\sum_{i=x}^r u_{TfTi} (\mathbf{R} - \mathbf{P}) > u_{TfTx-1} (\mathbf{P} - \mathbf{S}) + u_{TfTx} (\mathbf{T} - \mathbf{R}).$$

From this inequality it follows that for uniform ‘ $TfTx$ ’- $u$  distributions (i.e.,  $u_{TfT0} = u_{TfT1} = \dots = u_{TfTr}$ ), the distribution of the  $\pi$  values has a single peak at  $\pi_{TfTx}$  whereby  $x$  is the highest integer for which inequality  $(r+1-x)(\mathbf{R} - \mathbf{P}) > \mathbf{T} - \mathbf{S}$  is satisfied. As a consequence, direct effects can be obtained for uniform distributions by manipulating  $\mu$  if  $r(\mathbf{R} - \mathbf{P}) > \mathbf{T} - \mathbf{S}$  (i.e.,  $\pi_{TfT1} - \pi_{AllD} > 0$ ). The expectation that changes in conditions yielding increased  $x$ -values also result in increased execution of cooperation at the evolutionary equilibrium, was confirmed in a set of simulations (Spichtig, data not shown).

For  $rX$  sets, the following property should be noted. The response  $\rho_{ij}$  is the action sequence (of length  $r$ ) that strategy  $i$  triggers from strategy  $j$ , and  $\rho_i$  is the entire set of responses  $\rho_{ij}$  ( $j \in rX$ ) of strategy  $i$ . Given the comprehensiveness of  $rX$  sets it follows that – for arbitrary set  $\rho_i$  ( $i \in rX$ ) – the same number of respective responses is found for each of the  $2^r$  action sequences (i.e.,  $\rho_i$  and  $\rho_j$  ( $i \neq j$ ) are two permutations of the same set of sequences). The consequence is that, with uniform  $u$  distributions, the mean behaviors of mutants are not influenced by the strategy of the opponents. In that case, it can be inferred from the payoff dominance of  $\mathbf{D}$  over  $\mathbf{C}$  that  $AllD$  generates the absolute highest mean payoff from mutants ( $\pi_{AllD} > \pi_i$ ).

For the uniform distributions analyzed above, mean behaviors of mutants are not influenced by the strategy of the opponents. We refer to such  $u$  distributions with unbiased average mutant behaviors as symmetric and to alternative  $u$  distributions with biased average mutant behaviors as asymmetric. This distinction is useful because not only uniform  $u$  distributions of  $rX$  sets are symmetric. For example, any distribution with uniform  $u_i$  values for the conditional strategies is symmetric because the behavior

of unconditional strategies is not influenced by the opponent. In Appendix 3-1, we define the space of symmetric  $u$  distributions. Note, for both symmetric and asymmetric distributions, increasing the share of unconditional strategies tends to favor  $\pi_{AllD}$  as  $AllD$  expresses best response behavior to unconditional strategies. As outlined for the uniform distributions, direct effects can be excluded for all symmetric distributions. Hence, direct effects emerge only if strategies can trigger distinct mean mutant behaviors (i.e., the key characteristic of asymmetric distributions).

Symmetric  $u$  distributions for  $rX$  sets are a special case. The ‘ $TjTx$ ’ sets – as shown above – allow for direct effects, and they represent asymmetric distributions (the  $u_i$  values of ‘ $TjTx$ ’ sets are formed from  $u$  distributions of  $rX$  sets by setting the  $u_i$  values to zero for strategies outside the ‘ $TjTx$ ’ sets). It is apparent, for direct effects to occur, that average encounters with mutants should be inefficient for  $AllD$  but efficient for certain other strategies, i.e., mutants should tend to conditionally defect in interactions with  $AllD$  and should tend to conditionally cooperate in interactions with certain other strategies. Examples are distributions (such as ‘ $TjTx$ ’) for which mutants tend to express reciprocal behaviors (Trivers 1971; Axelrod and Hamilton 1981).

### 3.2 Simulations of $rX$ -populations

In order to gain insight in indirect effects, we performed simulations using the  $rX$  strategy sets  $\{r1, r2, r3, r4\}$  with uniform  $u$  distributions. As discussed in the previous subsection, direct effects are excluded with uniform  $u$  distributions. A set of simulations was performed with parameters  $\{T, S, R, K\}$  fixed to  $\{5, 0, 3, 0\}$ , while varying mutual defection payoff  $P = \{0.05, 0.3, 1\}$  and mutation rates  $\mu = \{0.0001, 0.001, 0.01, 0.1\}$ . For these parameter combinations, Table 3-1 shows whether populations evolve to an equilibrium or not (equilibrium conditions are described in Appendix 3-1). For all settings,  $\{r1, r2\}$ -populations (i.e., playing the one-round and the two-round game) evolve to equilibrium (Table 3-1). The table shows that for  $P = \{0.05, 0.3\}$ , no equilibrium is attained in the evolution of certain  $r3$ -populations and of certain  $r4$ -populations.

The equilibrium populations described in Table 3-1 are dominated by  $AllD$  (i.e.,  $f_{AllD} > f_i$  for  $i \neq AllD$ ) – this characteristic applies to all observed equilibrium populations in our study. Furthermore, all observed equilibrium strategy frequencies  $f_i$  are identical for both continuous and discrete-generation models. At equilibrium, dominance of  $AllD$  implies that the strategy also has fitness dominance. We do not find persistent indirect effects in the populations that do not reach equilibrium. Consequently, we do not find persistent indirect effects in the simulations.

For  $P = 1$ , the  $\{r1, r2, r3, r4\}$ -populations evolve to equilibrium for all mutation rates (Table 3-1). For rates  $\mu = \{0.001, 0.01, 0.1\}$ , Table 3-2a shows the average number of  $C$  executions per Prisoner’s Dilemma game ( $\frac{\bar{c}}{r}$ ) in these equilibrium populations. For each setting, these averages increase with mutation rates. The  $\frac{\bar{c}}{r}$ -values of  $r1$ -populations (Table 3-2a) are only slightly higher than the inflow of cooperator ( $c$ ) mutants ( $\sim 0.5\mu$ ). The execution of cooperation can thus be attributed to  $c$ -mutants. The table shows for each mutation rate that  $\frac{\bar{c}}{r}$ -values of  $\{r2, r3, r4\}$ -populations are approximately three times higher than those of  $r1$ . We attribute this difference to the

fact that sets  $\{r2, r3, r4\}$  contain conditional strategies. Table 3-2b shows that evolution to an equilibrium is found in simulations of  $r3$ -populations using the two background fitness values  $K = \{5, 20\}$ . Along  $K = \{0, 5, 20\}$  we find an increase in mean cooperation for each rate  $\mu$  (Table 3-2b).

The observed cooperation in the populations of Table 3-2 is maintained by mutation-selection balance because direct and indirect effects are absent. This interpretation of the  $\frac{\bar{C}}{r}$ -data is straightforward for the  $r1$ -populations. For the populations with repeated games, cooperation can be argued to be disadvantageous because ‘non-*AllD*’-individuals would increase their fitness by substituting their strategy for *AllD*. However, we emphasize that in several  $\{r3, r4\}$ -populations, non-defectors obtain above average fitness at equilibrium (Tables 3-1 and 3-2). The potential for the evolution of conditional behavior in repeated games ( $\{r2, r3, r4\}$ ) seems to reduce selection against cooperation (as cooperation levels are higher for these sets than in  $r1$ ; see Table 3-2a). As expected, a similar effect can be attributed to increasing background fitness  $K$  (Table 3-2b).

Table 3-1 shows that for the lowest mutual defection payoff ( $P = 0.05$ ),  $\{r3, r4\}$ -populations do not converge to equilibrium in the simulations with the two lowest mutation rates. For the intermediate  $P$ -value of Table 3-1, this phenomenon is also observed for  $r3$ -populations at the lowest rate and for  $r4$ -populations at the three lowest rates. With its 256 times smaller set size, the  $r3$ -populations are more convenient to study. This is why we mainly study non-equilibrium behavior in  $r3$ -populations.

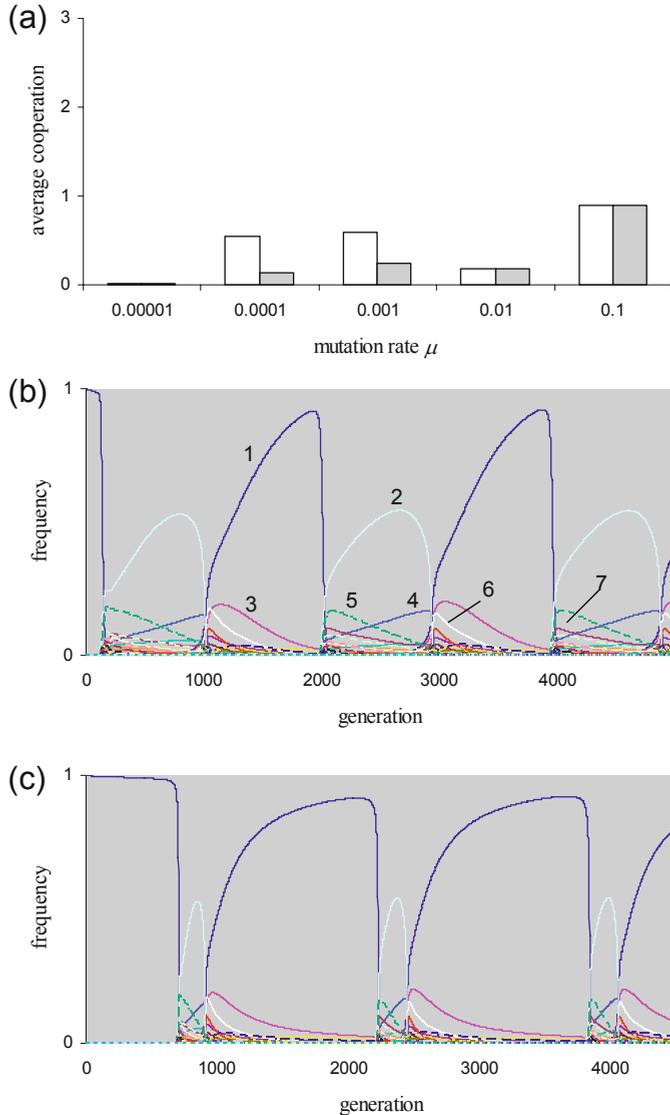
**Table 3-1** Observed type of dynamics in the final phases of simulations of four  $rX$  sets (top row) at three mutual defection payoffs (first column). For each  $\{rX, P\}$  combination, simulations were performed at the mutation rates  $\mu = \{0.0001, 0.001, 0.01, 0.1\}$ . In this alignment, letters  $\{n, e, E\}$  of the four digit strings represents the dynamics found in simulations at respective rate;  $n$ : non-equilibrium dynamics,  $e$ : equilibrium in which only defectors obtain above average fitness, and  $E$ : equilibrium in which non-defectors obtain above average fitness (for example,  $eeeE$  means that equilibrium is found at all four rates whereby non-defectors attain above average fitness only at rate  $\mu = 0.1$ ). Fixed parameters:  $T = 5$ ,  $S = 0$ ,  $R = 3$ ,  $K = 0$ .

|            | $r1$        | $r2$        | $r3$        | $r4$        |
|------------|-------------|-------------|-------------|-------------|
| $P = 0.05$ | <i>eeee</i> | <i>eeee</i> | <i>nmEE</i> | <i>nmEE</i> |
| $P = 0.3$  | <i>eeee</i> | <i>eeee</i> | <i>nEEE</i> | <i>nnnE</i> |
| $P = 1$    | <i>eeee</i> | <i>eeee</i> | <i>eeeE</i> | <i>eeeE</i> |

**Table 3-2** For equilibrium populations, the table shows the average amount of executed  $C$  actions per PD game ( $\frac{\bar{C}}{r}$ ) found at equilibrium, as a function of mutation rate. The top row gives mutation rates, first column the  $rX$  sets, and second column the background fitness values  $K$ .  $\frac{\bar{C}}{r}$ -values of equilibria in which non-defectors obtain above average fitness are given in italic. Fixed parameters:  $P = 1$ ,  $T = 5$ ,  $S = 0$ ,  $R = 3$ . (a) Varying strategy set for  $K = 0$ ; (b) varying background fitness for strategy set  $r3$ .

|          |      |          | $\mu = 0.001$ | $\mu = 0.01$ | $\mu = 0.1$   |
|----------|------|----------|---------------|--------------|---------------|
| <b>a</b> | $r1$ | $K = 0$  | 0.0005        | 0.0051       | 0.0577        |
|          | $r2$ | $K = 0$  | 0.0012        | 0.012        | 0.141         |
|          | $r3$ | $K = 0$  | 0.0014        | 0.0148       | <i>0.1844</i> |
|          | $r4$ | $K = 0$  | 0.0013        | 0.013        | <i>0.1712</i> |
| <b>b</b> | $r3$ | $K = 0$  | 0.0014        | 0.0148       | <i>0.1844</i> |
|          | $r3$ | $K = 5$  | 0.0038        | 0.0382       | <i>0.3294</i> |
|          | $r3$ | $K = 20$ | 0.0109        | 0.1002       | <i>0.43</i>   |

For  $P = 0.05$ , Fig. 3-1a shows the mean execution of cooperation per frPD game ( $\bar{C}$ ) along  $\mu = \{0.00001, 0.0001, 0.001, 0.01, 0.1\}$ . For the lowest rate and for the two highest rates, these means are sampled at equilibrium. As mentioned, the equilibrium



**Fig. 3-1** (a) Average execution of cooperation per game ( $\bar{C}$ ) as function of mutation rate in *r3*-simulations of *dgm*-populations (white) and of *cgm*-populations (gray). For rates  $\mu = \{10^{-5}, 0.01, 0.1\}$ , averages are sampled at equilibrium, and for rates  $\mu = \{0.0001, 0.001\}$ , averages are sampled over a cycle period (see panels (b) and (c)). Fixed parameters:  $P = 0.05$ ,  $T = 5$ ,  $S = 0$ ,  $R = 3$ ,  $K = 0$ . (b) For the conditions of panel (a), the frequency dynamics of an evolving *dgm*-population at rate  $\mu = 0.001$ . 1: *ddddddd* (*AllD*); 2: *cdddddd* (*TfT1*); 3: *ddddddc*; 4: *ddcdddd*; 5: *cdcdddd* (*TfT2*); 6: *dcdddd*; 7: *dccdddd*. (c) For the conditions of panel (a), the frequency dynamics of an evolving *cgm*-population at rate  $\mu = 0.001$ . Panel (b) and panel (c) use the same line code.

frequencies are not affected by the choice of the generation model (i.e., *dgm* or *cgm*). Hence, the  $\bar{C}$ -values are identical in Fig. 3-1a for each of these rates. After transient phases, the populations at rates  $\mu = \{0.0001, 0.001\}$  evolve in cycles. As an example, consider the strategy dynamics at rate  $\mu = 0.001$  in Fig. 3-1b for *dgm* and in Fig. 3-1c for *cgm*. Table 3-3 lists the strategies with  $\max(f_i) > 0.1$  during the cycles for these two figures. For mutation rates  $\mu = \{0.0001, 0.001\}$ , the  $\bar{C}$ -values in Fig. 3-1a are averaged over one cycle period. The  $\bar{C}$ -values are identical if populations are initialized with  $f_{AID} = 1$  and with a uniform frequency distribution. For both mutation rates, the averages  $\bar{C}$  are higher if sampled over *dgm*-cycles than if sampled over *cgm*-cycles (Fig. 3-1a). The figure also shows that for both models, the  $\bar{C}$ -values are higher in the cycling populations than for the equilibrium populations at  $\mu = 10^{-5}$ . The  $\bar{C}$ -values are higher than the equilibrium-values found at the higher rate  $\mu = 0.01$  for the *dgm* at rates  $\mu = \{0.0001, 0.001\}$  and for the *cgm* at rate  $\mu = 0.001$  (Fig. 3-1a). Consequently, for both types of generation models an optimum in  $\bar{C}$  exists within the interval  $10^{-5} < \mu < 0.01$ .

For the two mutation rates  $\mu = \{0.0001, 0.001\}$ , we tested the sensitivity of the cycling dynamics in *dgm*-populations to the choice of background fitness  $K$ . The populations show cycling dynamics if  $K \leq \{45, 3\}$  ( $\rightarrow \mu = \{0.0001, 0.001\}$ ) and evolve to equilibrium for higher  $K$ -values. The *r3*-populations showing non-equilibrium dynamics in our simulations evolve in cycles (and show periodic indirect effects).

The strategy dynamics in Fig. 3-1b,c resemble those in the corresponding simulations with the lower mutation rate  $\mu = 0.0001$ . All four cycles show (as in Fig. 3-1b,c) alterations of phases with dominance of *AID* followed by phases with dominance of *TfT1* (*cdddddd*). As can be inferred from these dynamics, *AID* respectively *TfT1* have the highest fitness when invading the populations. Consequently, these populations express periodic indirect effects. In Fig. 3-2, we give behavioral statistics from the simulation of Fig. 3-1b. Fig. 3-2a shows the dynamics of the mean number of executed  $C$  actions for each round of the game ( $\bar{C}_i, i = \{1, 2, 3\}$ ). Cooperation is more intensively executed during *TfT1* phases, especially in round 1 (Fig. 3-2a). The relatively longer *TfT1* phase durations in the *dgm*-populations (compare Fig. 3-1b with 3-1c) explain that  $\bar{C}$ -values are higher in *dgm*-populations than in corresponding *cgm*-populations (Fig. 3-1a at  $\mu = \{0.0001, 0.001\}$ ).

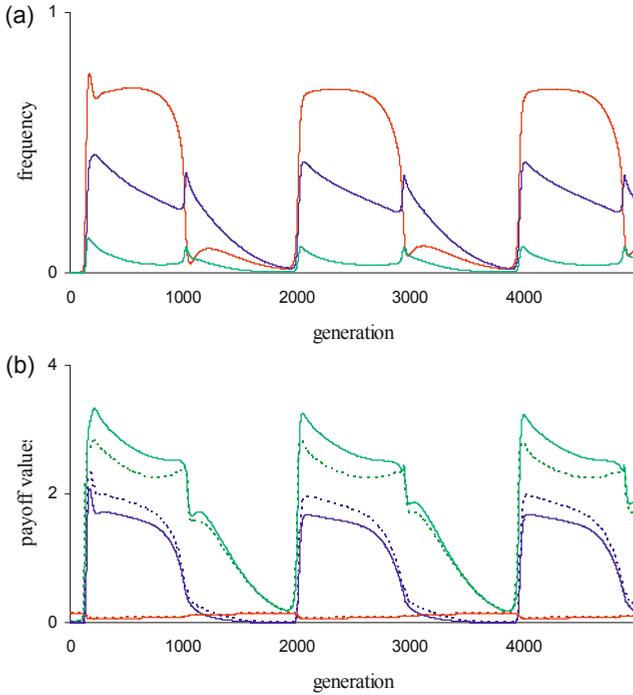
**Table 3-3** List of strategies that obtain peak frequencies higher than 0.1 ( $\max(f_i) > 0.1$ ) within the cycle phases of the *dgm*-dynamics depicted in Fig. 3-1b. Code representation (conventional name in brackets) of the strategies is given in the second column. The peak frequency within the cycle phases is given in the third column. The peak frequency within the cycle phases of the *cgm*-dynamics of Fig. 3-1c is given in the fourth column.

|   | Strategy                       | $\max(f_i)$ | $\max(f_i)$ |
|---|--------------------------------|-------------|-------------|
| 1 | <i>dddddd</i> ( <i>AID</i> )   | 0.921       | 0.921       |
| 2 | <i>cdddddd</i> ( <i>TfT1</i> ) | 0.544       | 0.544       |
| 3 | <i>dddddc</i>                  | 0.201       | 0.201       |
| 4 | <i>dcdddd</i>                  | 0.168       | 0.168       |
| 5 | <i>cdcdddd</i> ( <i>TfT2</i> ) | 0.166       | 0.165       |
| 6 | <i>dcdddd</i>                  | 0.159       | 0.157       |
| 7 | <i>dccdddd</i>                 | 0.103       | 0.102       |

For the three payoffs  $\mathbf{P}$ ,  $\mathbf{T}$ , and  $\mathbf{R}$ , Fig. 3-2b shows the dynamics of the mean payoff values per frPD game (i.e.,  $\overline{p_P}$ ,  $\overline{p_T}$ , and  $\overline{p_R}$ ). Steep increases in the generation of  $\mathbf{T}$ -payoffs (Fig. 3-2b) mark the onset of invasions by  $TfT1$  (Fig. 3-1b). Defectors like  $AllD$  generate this payoff in the first round when interacting with  $TfT1$  and defectors are the dominant opponents of this strategy at the onset of invasions (Fig. 3-1b). The increase in the generation of  $\mathbf{T}$ -payoffs is therefore partly explained by defectors triggering this payoff from  $TfT1$ . For  $TfT1$ , these first round interactions seem disadvantageous, but this disadvantage is evidently compensated because  $TfT1$  invades.

Fig. 3-2b additionally shows the dynamics of expected average payoff values generated per game from receiving payoff  $\mathbf{P}$ ,  $\mathbf{T}$ , or  $\mathbf{R}$  (i.e.,  $P \sum_{i \in \{1,2,3\}} (1-\overline{C}_i)^2$ ,  $T \sum_{i \in \{1,2,3\}} (1-\overline{C}_i)\overline{C}_i$ , and  $R \sum_{i \in \{1,2,3\}} \overline{C}_i^2$ ). For payoff  $\mathbf{T}$ , the observed value is higher than the expected value

(Fig. 3-2b) over the dominance phase of  $TfT1$  (Fig. 3-1b). These differences between observed and expected values are caused by the conditional behaviors in rounds 2 and 3. Hence, we propose that the invasions of  $TfT1$  are fueled by triggering  $\mathbf{T}$ -payoffs in



**Fig. 3-2** Both panels show behavioral statistics of the evolving populations in Fig. 3-1b. (a) Average execution of cooperation in round 1 ( $\overline{C}_1$ : red), in round 2 ( $\overline{C}_2$ : blue), and in round 3 ( $\overline{C}_3$ : green) as function of time. (b) Average payoff generated per game with payoff  $\mathbf{P}$  (red), payoff  $\mathbf{T}$  (green), and payoff  $\mathbf{R}$  (blue) as function of time. Dashed lines represent respective expected values (i.e.,

$P \sum_{i \in \{1,2,3\}} (1-\overline{C}_i)^2$ ,  $T \sum_{i \in \{1,2,3\}} (1-\overline{C}_i)\overline{C}_i$ , and  $R \sum_{i \in \{1,2,3\}} \overline{C}_i^2$ ) and solid lines represent respective observed values (i.e.,  $\overline{p_P}$ ,  $\overline{p_T}$ , and  $\overline{p_R}$ ).

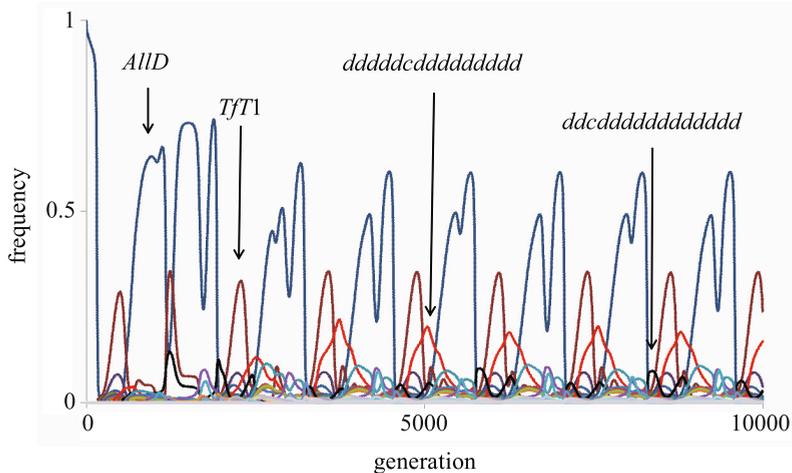
these rounds. At the onset of invasions, *AllD* is the dominant strategy (Fig. 3-1b) and defection is the predominant behavior (Fig. 3-2). Defectors (in contrast to non-defectors) are not penalized when interacting with *AllD* and they can therefore be expected to perform better than other strategies in *AllD*-dominated populations. The strategy *TfT1* generates *T*-payoffs from the twelve defectors  $\{ddcd\dots, dddd.c.\}$ . Game interactions between these defectors and *TfT1* indeed significantly contribute (data not shown) to the increases of  $\overline{p_T}$  (Fig. 3-2b).

In Appendix 3-1, we derive the invasion condition for a single *TfT1*-individual in a population state with full defection. We find that such invasion occurs if the combined frequency of defectors  $\{ddcd\dots, dddd.c.\}$  exceeds  $(P - S)/(T - P)$  ( $\sim 0.01$  in the simulation of Fig. 3-1b). This condition is fulfilled over the entire cycle period in Fig. 3-1b, but the population state deviates from full defection due to mutation. In this state, *AllD* obtains the highest benefit from interactions with mutants (i.e.,  $\mu(\pi_{AllD} - \pi_{TfT1}) > 0$ ). Thus, the invasion conditions in the simulations should be more stringent than those derived in Appendix 3-1. Before the onset of the invasions, the population does converge towards a state of full defection (Fig. 3-2) and thus towards the conditions underlying the analysis in Appendix 3-1. In our opinion, the invasions of *TfT1* in the simulations are fueled by interactions with defectors  $\{ddcd\dots, dddd.c.\}$ , just like in the analysis. That *AllD* subsequently regains dominance, thereby closing the cycle (Fig. 3-1b,c), is in line with the expectations from the selection dynamics of evolutionary frPD games (Nachbar 1992; Cressman 1996).

In Tables 3-1 and 3-2, we mark the equilibria (*E* in Table 3-1 and italic numbers in Table 3-2) in which non-defectors have above average fitness. The strategy *TfT1* has the highest fitness among the non-defectors in these equilibria. Furthermore, these equilibria emerge at the higher mutation rates (Tables 3-1 and 3-2) possibly because mutation benefits *TfT1* (e.g., by generating defectors  $\{ddcd\dots, dddd.c.\}$  opponents) in these equilibria. However, invasion by this strategy is prevented also because *AllD* is the strategy that benefits most from interactions with mutants ( $\mu(\pi_{AllD} - \pi_{TfT1}) > 0$ ).

The *r4*-simulations are more computation-intensive than the *r3*-simulations and we restricted these simulations to  $10^4$  generations due to constraints on computation time. Consequently, the data obtained do not allow definitive conclusions on the nature of non-equilibrium *r4*-dynamics. Over the simulation periods, chaotic dynamics occurs for the *r4*-populations with non-equilibrium dynamics in Table 3-1. For example, in Fig. 3-3 the *r4*-frequency dynamics at  $\{P, \mu\} = \{0.3, 0.01\}$  exhibits a transient period of  $\sim 2000$  generations, after which alternations of dominance by strategies  $\{AllD, ddcddddddddddd, TfT1, ddddcdcccccccc\}$  emerge. The population therefore expresses periodic indirect effects. As in Fig. 3-1b,c, strategies *AllD* and *TfT1* in Fig. 3-3 become periodically dominant with dominance phases of strategies  $\{AllD, TfT1, ddddcdcccccccc\}$  that have fairly regular phase lengths (Fig. 3-3).

For  $\{r1, r2, r3, r4\}$ , Table 3-1 shows that all populations evolve to equilibrium in the two smallest sets  $\{r1, r2\}$ , and non-equilibrium dynamics occurs more frequently when going from set *r3* to set *r4* (e.g., the *r3*-population evolves to equilibrium under the conditions of Fig. 3-3). We interpret this observation as an indication that increasing the number of rounds (*r*) increases the parameter range for which periodic indirect effects emerge. This interpretation meets our intuition because strategy *TfT1* generates one *T*-payoff from *r2*-defector *ddc*, two *T*-payoffs from *r3*-defectors *ddcdc.*, and three *T*-payoffs from *r4*-defectors *ddcdc.d...c....*



**Fig. 3-3** The frequency dynamics of an  $r4$ -simulation at  $\{P, \mu\} = \{0.3, 0.01\}$ ; other conditions as in Fig. 3-1. After a transient period of ca. 2000 generations, four strategies periodically reach dominance (i.e., highest frequency) in the following sequence:  $\{AllD$  (dark blue line),  $ddcddddddddddd$  (black line),  $TfT1$  (brown line),  $dddddcdddddddd$  (red line) $\}$ . Arrows indicate the line for these strategies.

## 4 Discussion

Whereas empirical ecologists typically observe wide behavioral variation, theoretical ecologists tend to ignore or minimize behavioral variation in their models in order to make their analyses tractable. In this paper, we provide a method to analyze effects of behavioral variation on evolutionary dynamics, and apply it to the evolution of cooperation. We present a model in which behavioral variation is on the one hand subject to a restriction because probabilistic strategies are excluded, yet on the other hand comprehensive because all deterministic strategies are taken into account [see Axelrod (1984), Lindgren (1992), Hauert and Schuster (1997) and van Veelen and Garcia (2011) for earlier studies of evolutionary repeated games with comprehensive strategy sets]. We first discuss the method of analysis, and then when and how behavioral variation affects the evolution of cooperation.

### 4.1 Direct and indirect effects of mutation-generated variation

We consider direct and indirect effects of mutation in frequency dependent selection environment. Direct effects are fitness effects that emerge from (game) interactions with mutants and indirect effects are fitness effects that emerge from interactions with descendants of mutants. Direct effects are mainly determined by the mutation regime (as reflected in the  $u$  distribution of mutants over all possible strategies). Indirect effects are additionally influenced by selection on the progeny of mutants. Both in direct and indirect effects invading strategies depend on the presence of other strategies [see van Veelen (2012) for analysis of invasions that depend on other strategies].

In our analysis of direct and indirect effects of mutation in asexual populations, we take advantage of simplifying cladistics: each mutant is the founder of a clade whereby the clade constitutes the clonal descendants that faithfully inherit the genome of the

mutant. Direct effects are caused by interactions with founders and indirect effects are caused by interactions with descendants. Would we have modeled sexual reproduction, then this would complicate the cladistics as clades crossover thereby generating new behavioral variants in another way than by mutation. Cultural transmission, however, could have similarly simple cladistics as our asexual model: a system with innovators and imitators could to some degree be analogous to our system with mutants and descendants (non-mutants).

To analyze direct effects for large strategy sets, we assume constant  $u$  distributions resulting in constant returns ( $\pi_i$ ) from the interactions with mutants. In nature,  $u$  distributions are probably variable: for example, the  $u$  distribution is variable if mutation swaps single code positions rather than modifies entire codes/strategies (as in our study). Variable  $u$  distributions would complicate the analysis of direct effects because the population composition has to be considered (whereby genotype  $\times$  mutation interactions, i.e., genotypes differ in their propensity to mutate, would further complicate this analysis).

We also assume that fitness differences concern differences in fertility. Would we have considered fitness differences in viability, this would complicate the determination of the  $\pi_i$ -values. For example, it is not clear to us whether  $\pi_i$ -values are still constants with constant  $u$  distributions. The fraction of mutants would definitely deviate from  $\mu$ . However, would we consider differences in viability (rather than fertility), then we cannot think of a reason why our qualitative findings with respect to the ‘ $TfTx$ ’ sets, the  $rX$  sets, and the symmetric  $u$  distributions would change.

Indirect effects require sufficiently strong frequency-dependent selection: they only emerge when the strength of selection is high. Indeed, indirect effects do not emerge in our simulations above certain values of background fitness  $K$ . In order to demonstrate indirect effects our method is best applied to a system with a unique attractor in the selection environment in absence of mutation. This attractor in turn is best represented by a strategy that benefits most from interactions with the mutants, because, otherwise, direct effects could blur indirect effects.

## 4.2 Direct effects as a mechanism promoting the evolution of cooperation

We use an evolutionary game version of the finitely repeated Prisoner’s Dilemma (frPD), because it has a unique attractor in absence of mutation: unconditional defector ( $AllD$ ) (Cressman 1996). This provides a straightforward criterion to test what happens when mutation is included: mutation has significant effects whenever at least one strategy persistently or periodically achieves a higher fitness than that of  $AllD$ .

We show conditions for direct effects when mutation produces ‘ $TfTx$ ’ strategies (Nachbar 1992), which permit the evolution of cooperation. We further show that the evolution of cooperation through direct effects are excluded for  $rX$  strategy sets (Cressman 1992) with a uniform  $u$  distribution. For the latter analysis, we define a class of  $u$  distributions, the symmetric  $u$  distributions (see Appendix 3-1: uniform  $u$  distributions of  $rX$  sets are examples of symmetric  $u$  distributions), for which the average (conditional) behavior of mutants is independent of the opponent strategy. Whenever the opponent strategy does not influence the average behavior of mutants, then  $AllD$  is the opponent strategy that receives the highest payoff from the interactions with mutants. Hence, direct effects are excluded for symmetric  $u$  distributions.

The ‘*TfTx*’ sets exemplify how to create mutation regimes that can elicit direct effects resulting in the evolution of cooperation. The *TfTx* strategies are contained in the corresponding *rX* set and the *rX* sets are subsets of the (infinite) space of probabilistic strategies adjusted to frPD games. We introduce the symmetric *u* distributions because they constitute a boundary in the *u* distribution space: they separate the subspace where mutants cooperate most often with *AllD* from the subspace where mutants cooperate most often with another strategy. Direct effects can only emerge in the latter subspace, depending on the conditions of the frPD game and the mutation rate. Note that the emergence of direct effects does not imply that the population evolves to a state where cooperation is amply executed (e.g., if the strategy benefiting most from the mutants is a defector).

Direct effects that result in the evolution of cooperation emerge only in a fraction of the *u* distribution space. We can only speculate about the size of this fraction. Even if this fraction is tiny, it may be important for the evolution of cooperation if *u* distributions in natural systems would fall into this category. Nevertheless, the empirical evidence for direct reciprocity (in general) is scarce (Clutton-Brock 2009), let alone evidence for mutation regimes. In any case, for evolutionary games with a clear attractor strategy like *AllD* for evolutionary frPD games, we conjecture that this strategy is most likely to benefit most from the mutants. Consequently, the evolution of cooperation by direct effects is possible but we predict it is not very likely in general.

The study of McNamara et al. (2004) inspired our definition of direct effects. In our view, the evolution of cooperation in the studies of McNamara et al. (2008), Eriksson and Lindgren (2005), and Traulsen et al. (2009) is explained by direct effects. The strategy sets used in these studies are only part of a much larger set of (deterministic) strategies, just like the ‘*TfTx*’ sets in relation to the *rX* sets. Furthermore, if their mutation regimes were replaced by regimes comprising broader strategy sets, then we would expect that the unconditional defectors in their evolutionary games benefit most from interactions with mutants. Hence, we think that the mutation-induced promotion of the evolution of cooperation, as described by McNamara et al. (2004, 2008), Eriksson and Lindgren (2005), and Traulsen et al. (2009), is a rather special outcome. If, however, behavioral variation is not caused by mutation, but culturally-inherited, then cooperation may evolve under a wider set of conditions if it is true that humans choose from ‘*TfTx*’ strategy sets and disregard most *rX* strategies.

### 4.3 Indirect effects as a mechanism promoting the evolution of cooperation

In simulations of *rX* sets with uniform *u* distributions (i.e., a condition without direct effects), we observe periodic indirect effects (Fig. 3-1b,c and 3-3). We find several conditions where such effects emerge (Table 3-1) at intermediate mutation rates (Fig. 3-1a). The periodic indirect effects all show a similar pattern (Fig. 3-1b,c and 3-3): a population dominated by *AllD* is invaded by strategy *TfT1* and vice versa, giving rise to cycles of alternating dominance. We suggest that the invasion of *TfT1* is due to a group of defectors defined by  $\{ddcd\dots, dddd.c.\}$  (see also Appendix 3-1C). The behavior of these defectors towards other defectors is similar to the dominant behavior in *AllD*-dominated populations: play defect in all rounds of the game. Therefore, these defectors are less vulnerable to the (*AllD*-influenced) selection in such populations and decrease less (due to mutation-selection balance) than other *rX* strategies. *TfT1* exploits

the behavioral deviations that certain defectors have from the behavior of *AllD* (see Appendix 3-1C). As a consequence, the execution of cooperation increases during *TfT*-invasions (Fig. 3-2a) and the average execution of cooperation can exceed the execution value expected from a mutation-selection balance (Fig. 3-1a).

We only found periodic indirect effects. We suspect that – for evolutionary frPD games (without direct effects) – persistent indirect effects emerge only for special parameter regions, if they exist at all. This is because they require a strong enough effect of the *AllD* individuals on the fitness of others already before this strategy fully achieves fitness dominance.

The periodic indirect effects observed in our study resulted in periodic increases in the execution of the action ‘cooperation’ (Fig. 3-2a). But one may ask whether this increase constitutes co-operation in the sense of individuals mutually helping each other. If *TfT* invades a population otherwise composed by *AllD* then after the first round the *TfT*-players cooperate only with other *TfT*-players. Cooperation therefore mostly takes place among individuals with the same phenotype (i.e., *TfT* players) and the mutual cooperation payoff is generated more often than expected. Such positive assortment is known as a fundamental principle for the evolution of cooperation (Nowak 2006). In our simulations, we find the opposite: decisive executions of cooperation take place between individuals with different phenotypes, i.e., *TfT* and defectors  $\{ddcd\dots, dddd.c.\}$ , and the mutual cooperation payoff is generated less often than expected (Fig. 3-2b). Furthermore, in the interactions between *TfT* and these defectors it is not beneficial for the defectors to stick to their strategy, as they would fare better by playing unconditional defection. On the other hand, it is typical for the evolution of cooperation that the average payoff increases, as observed during *TfT* invasions (Fig. 3-2b).

Strategy *TfT* invades in the large set of strategies that invade in the wake of *TfT* invasions (Fig. 3-1b,c and 3-3). This conditional cooperator might play a more pronounced role in frPD games with more than four rounds (as increasing  $r$  promotes the performance of *TfT*). Unfortunately, we cannot check this prediction because the sheer size of the corresponding  $rX$  sets (e.g., 2147483648 strategies in a game with five rounds) makes running simulations with five or more rounds unfeasible.

#### 4.4 Conclusions

By analyzing direct effects, we explain the phenomenon of mutation-promoted evolution of cooperation, described earlier by McNamara et al. (2004, 2008), Eriksson and Lindgren (2005) and Traulsen et al. (2009). However, we argue that this phenomenon is probably a rather special case: the strategy favored by selection without mutation – *AllD* in our study – is most likely also the strategy with the highest benefit from interactions with mutants (direct effect). In such cases, cooperation can still evolve as an indirect effect of mutation-generated variation under a limited set of conditions. The resulting cooperation dynamics, however, shows exploitation of cooperative acts rather than mutual cooperation.

The study of McNamara et al. (2004) is seminal in highlighting the importance of behavioral variation in evolutionary dynamics. Theoreticians have avoided this topic because behavioral variation complicates model analysis. To facilitate such analysis, our method to separate direct and indirect effects of behavioral variation is a useful

approach to assess if and when behavioral variation is important in evolutionary dynamics.

### Appendix 3-1

#### A: The symmetric $u$ distributions

A  $u$  distribution is symmetric if the average conditional game behavior of mutants is independent of the game behavior that the opponent expresses. That is, given a mutant sampled from this distribution, if this mutant expresses conditional behavior in arbitrary round  $x$  then the probability that this mutant defects in round  $x$  is independent of the action sequence the opponent played in the previous rounds. In the following, we demonstrate the symmetric  $u$  distributions.

Game behavior can be conditional from round 2 onwards. Strategies carrying  $dd$  or  $cc$  at code positions 2 and 3 express unconditional behavior in round 2. Strategies carrying  $dc$  or  $cd$  express conditional behavior in this round. Similarly, strategies carrying  $dddd$  or  $cccc$  at positions 4 to 7 express unconditional behavior in round 3. The strategies with a different code at these positions express conditional behavior in round 3. Also for the higher rounds, two codes determine unconditional behavior and the remaining codes determine unconditional behavior in that round.

Constant  $u$  distributions are symmetric if each round behavior mutates according to the following rule. In round  $x$ , the mutant expresses unconditional defection with probability  $u_{x1}$ , unconditional cooperation with probability  $u_{x2}$ , and conditional behavior with probability  $u_{x3}$  ( $= 1 - u_{x1} - u_{x2}$ ). In case of the latter mutation event, the probability that the mutant carries either code determining conditional round  $x$  behavior requires the following property: the chance that the mutant defects (cooperates) in this round is independent of the previous action sequence played by the opponent. This condition is given if either of these codes is carried by the mutant with equal probability.

Variable  $u$  distributions are symmetric if each round behavior mutates according to the following rule. Mutation treats parental genotypes with unconditional round  $x$  defection, with unconditional round  $x$  cooperation, and with conditional round  $x$  behavior differently. Unconditional round  $x$  defection (cooperation) is faithfully inherited with probability  $1 - u_{xD}$  ( $1 - u_{xC}$ ) and mutates otherwise. Conditional round  $x$  behavior always mutates. Dependent on the parental genotype, we therefore find three forms of mutation events. For each form, mutation proceeds analogously as described for constant symmetric  $u$  distributions (whereby the three probabilities  $\{u_{x1}, u_{x2}, u_{x3}\}$  can be distinct between the forms).

#### B: Equilibrium conditions

We assess equilibrium in  $\{r1, r2, r3\}$ -populations as reached if all differences ( $f'_i - f_i$ ) in the  $dgm$  are less than  $10^{-8}$  between consecutive generations. Note, the corresponding  $cgm$ -populations evolve to the same frequency distributions for all observed equilibria. The approached equilibria are identical for initial  $r1$ -populations  $\{f_D = 0, f_D = 0.5, f_D = 1\}$ . The equilibrium  $\{r2, r3\}$ -populations are identical if initiated with  $f_{AllD} = 1$  or

if initiated with uniform frequency distributions. Only the *dgm*-model is implemented to simulate *r*4-populations and all *r*4-populations are initiated with  $f_{AID} = 1$ . *r*4-equilibrium is assumed if all differences  $f_i' - f_i$  decrease over consecutive generations for a period of 2000 generations.

### C: Invasion condition of *TfT1* in *r*3-defector populations

We study the fitness of a *TfT1* individual in populations otherwise composed of defectors (full defection). Defectors B are the strategies  $\{ddcd\dots, dddd.c.\}$ . Defectors A are the remaining defectors of *r*3. Strategy *TfT1* earns payoff  $(S + 2P)$  from interactions with defector A individuals, payoff  $(S + 2T)$  from interactions with defector B subgroup *ddcd.c.*, and payoff  $(S + P + T)$  from interactions with the other defector B subgroup. The combined frequency of defectors B strategies is defined as  $f_B$ . We calculate the most stringent condition for invasion of *TfT1*, i.e., assuming that the defectors *ddcd.c.*, from which *TfT1* generates the higher payoff, are absent. Then, *TfT1* individuals should obtain above average payoffs if  $f_B (S + P + T) + (1 - f_B) (S + 2P) > 3P$ . Consequently, negligibly small  $f_{TfT1}$ -values increase if  $f_B > (P - S)/(T - P)$ .

## References

- Axelrod R (1984) The evolution of cooperation. Basic Books. New York
- Camerer CF (2003) Behavioral Game Theory: Experiments in Strategic Interaction, Princeton University Press
- Clutton-Brock T (2009) Cooperation between non-kin in animal Societies. Nature 462:51-57
- Cressman R (1996) Evolutionary stability in the finitely repeated prisoner's dilemma game. J Econ Theor 68:234-248
- Eriksson A, Lindgren K (2005) Cooperation driven by mutations in multi-person Prisoner's Dilemma. J Theor Biol 232:399-409
- Fischbacher U, Gächter S, Fehr E (2001) Are people conditionally cooperative? Evidence from a public goods experiment. Econ Letters 71:397-404
- Fudenberg D, Tirole J (1991) Game Theory. Cambridge MA: MIT Press
- Hauert Ch, Schuster HG (1997) Effects of increasing the number of players and memory steps in the iterated Prisoner's Dilemma, a numerical approach. Proc R Soc Lond B 264:513-519
- Henrich J, McElreath R, Barr A, Ensminger J, Barrett C, Bolyanatz A, Cardenas JC, Gurven M, Gwako E, Henrich N, Lesorogol C, Marlowe F, Tracer D & Ziker J (2006) Costly punishment across human societies. Science 312:1767-1770
- Herrmann B, Thöni C, Gächter S (2008) Antisocial punishment across societies. Science 319:1362-1367
- Kurzban R, Houser D (2005) Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations. PNAS 102:1803-1807
- Lindgren K (1992) Evolutionary phenomena in simple dynamics. In: Artificial Life II, C. Langton et al. (Eds.), pp. 295-312. Addison-Wesley, Redwood City
- Maynard Smith J (1982) Evolution and the theory of games. Cambridge Univ. Press, Cambridge
- McNamara JM, Barta Z, Houston AI (2004) Variation in behaviour promotes cooperation in the prisoner's dilemma game. Nature 428:745-748
- McNamara JM, Barta Z, Fromhage L, Houston AI (2008) The coevolution of choosiness and cooperation. Nature 451:189-192
- McNamara JM, Leimar O (2010) Variation and the response to variation as a basis for successful cooperation. Phil Trans R Soc B 365:2627-2633
- Nachbar JH (1992) Evolution in the finitely repeated prisoner's dilemma. J Econ Beh Org 19:307-326
- Nowak MA (2006) Five rules for the evolution of cooperation. Science 314:1560-1565
- Page KM, Nowak MA (2002) Unifying evolutionary dynamics. J Theor Biol 219:93-98
- Traulsen A, Hauert Ch, De Silva H, Nowak MA, Sigmund K (2009) Exploration dynamics in evolutionary games. PNAS 106:709-712

- Wedekind C, Milinski M (1996) Human cooperation in the simultaneous and the alternating Prisoner's Dilemma: Pavlov versus Generous Tit-for-Tat. PNAS 93:2686-2689
- West SA, Griffin AS, Gardner A (2007) Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. J Evol Biol 20:415-432
- Willensdorfer M, Nowak MA (2005) Mutation in evolutionary games can increase average fitness at equilibrium. J Theor Biol 237:355-362
- van Veelen M (2012) Robustness against indirect invasions. Games Econ Behav 74:382-393