

The evolution of payoff matrices: providing incentives to cooperate

Erol Akçay^{1,*} and Joan Roughgarden²

¹National Institute for Mathematical and Biological Synthesis (NIMBioS), 1534 White Avenue, Knoxville, TN 37996, USA

²Department of Biology, Stanford University, 371 Serra Mall, Stanford, CA 94305, USA

Most of the work in evolutionary game theory starts with a model of a social situation that gives rise to a particular payoff matrix and analyses how behaviour evolves through natural selection. Here, we invert this approach and ask, given a model of how individuals behave, how the payoff matrix will evolve through natural selection. In particular, we ask whether a prisoner's dilemma game is stable against invasions by mutant genotypes that alter the payoffs. To answer this question, we develop a two-tiered framework with goal-oriented dynamics at the behavioural time scale and a diploid population genetic model at the evolutionary time scale. Our results are two-fold: first, we show that the prisoner's dilemma is subject to invasions by mutants that provide incentives for cooperation to their partners, and that the resulting game is a coordination game similar to the hawk–dove game. Second, we find that for a large class of mutants and symmetric games, a stable genetic polymorphism will exist in the locus determining the payoff matrix, resulting in a complex pattern of behavioural diversity in the population. Our results highlight the importance of considering the evolution of payoff matrices to understand the evolution of animal social systems.

Keywords: evolutionary game theory; prisoner's dilemma; hawk–dove game; behavioural dynamics; two-tiered model

1. INTRODUCTION

Evolutionary game theory (EGT) is one of the fundamental tools to study how behaviour and traits of organisms evolve by natural selection. An *evolutionary game* is defined by different genetical strategies and their fitness (i.e. reproductive output) when they interact with each other. The genotypes increase or decrease in frequency according to their fitness given the frequencies of other strategies in the population. This process frequently (but not always) leads to an evolutionarily stable strategy (ESS), which is a strategy that, when fixed in the population, cannot be invaded by alternative strategies.

Earlier EGT models in biology tended to assume that the genetical strategies correspond to actual behaviours [1], or simple conditional rules that prescribe a certain behaviour given the state of the individual and the interaction (e.g. the tit-for-tat strategy [2]). More recent work has focused on interactions where individuals' behaviour is not directly determined by their genes, but instead reflect the outcome of a dynamical process where the players respond to each other according to proximate mechanisms that prescribe their behaviour [3–7]. In particular, Roughgarden [8] calls for an explicitly two-tiered conception of behavioural evolution: the first tier describes the dynamics of behaviour within the time scale of an interaction where individuals can adjust their actions in response to the context and the behaviours of others. The second tier, on the other

hand, is defined by the usual evolutionary game, with the distinction that the fitness values of the individuals are determined by the result of the behavioural dynamics in the first tier.

A two-tiered conception of behavioural evolution calls for more explicit models of proximate mechanisms of behaviour, instead of relying on implicit assumptions and ad hoc interpretations of ESS outcomes. Furthermore, the introduction of behavioural dynamics opens the door to new questions that have been overlooked previously, but have important evolutionary consequences. The first such question is what the dynamics at the behavioural tier look like. The two-tiered approach allows dynamics where individuals may coordinate their actions and act in concert with each other even if their fitness interests are not aligned completely [6,9], and explicitly model the evolution of other-regarding motivations [10].

In this paper, we are concerned with a second, complementary question; namely how the payoffs from the social interaction themselves can evolve. To be more precise, we make a clear distinction between the 'behavioural game', which consists of the observable actions and material payoffs to the individuals, and the 'evolutionary game', which is played between genetical strategies at the population level. The same behavioural game can give rise to many different evolutionary games, depending on how behaviour is determined and whether individuals interact with each other repeatedly, etc. Most work in EGT is concerned with taking a given behavioural game and transforming it into different evolutionary games by adding different rules for decision-making. In contrast, we fix the decision-making rule, and look instead at how the payoffs from the behavioural game can evolve. To our knowledge, only one previous study [11] has taken up this question before.

* Author for correspondence (erol@nimbios.org).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2010.2105> or via <http://rspb.royalsocietypublishing.org>.

We illustrate our argument with a simple example below; a more general treatment is provided in the following sections.

(a) A motivating example

Consider two male birds that defend adjacent territories. Each one can either fight (F) or make peace (MP) with the other [12]. Fighting is costly in terms of time and energy, and carries the risk of injury. Making peace avoids these costs, but when a bird tries to make peace unilaterally, it loses its territory to its fighting opponent. This description of the possible behaviours and their material consequences constitutes the behavioural game, which in this instance has the familiar structure of the prisoner’s dilemma:

		male 2		
		make peace (MP)	fight (F)	
	MP	3, 3	0, 5	(1.1)
male 1	F	5, 0	1, 1.	

The payoffs to the first and second males are given by the first and second number in each cell, respectively. They stand for the material costs and benefits individuals experience as the result of the social interaction—for example, the territory area of an individual after the interaction minus any effort spent on fighting.

When confronted with such a behavioural game, evolutionary game theorists have to make several decisions to translate it into an evolutionary game and ask which behaviours will emerge from natural selection. The most basic EGT assumptions are to take a very large population, match every individual at random to play the behavioural game only once and let the action of the individual be determined by its genetic locus. Under these assumptions, the evolutionary game looks exactly the same as the behavioural game, and it is easy to see that the alleles for fighting will increase in the population, as they have a higher fitness.

On the other hand, if the behavioural game is being played repeatedly, there can be other types of alleles that prescribe conditional behaviour, such as the tit-for-tat strategy [2,12], where individuals behave aggressively against fighting neighbours but not peaceful ones (e.g. [13]). In this case, the strategies in the evolutionary game include different conditional strategies. The fitness of each strategy will depend also on factors such as how many rounds of the game are played, whether the game is played against the same opponents, and so on. Thus, one can construct many evolutionary games from a single repeated behavioural game by adding new decision rules as possible evolutionary strategies, so that even if the behavioural game is a prisoner’s dilemma, the evolutionary game rarely is.

Another way the evolutionary game can be transformed is to not change the decision-making rule, but alter the payoffs from the behavioural game itself. For example, take a population playing game (1.1), where initially all individuals fight, and suppose that one of the males (say male 2) carries a mutation that reduces the level of testosterone in circulation during territorial contests. Suppose that due to this mutation, male 2 ‘pulls his punches’ when fighting and displays reduced aggressiveness outside its territory. Label this modified fighting strategy F*. As a consequence, the mutant male 2 would leave a larger share of the territory to

male 1 even when the latter is not fighting. In this case, the payoff matrix of the behavioural game between the mutant male 2 and resident male 1 might look like this:

		male 2		
		MP	F*	
	MP	3, 3	1.5, 3.5	(1.2)
male 1	F	5, 0	1, 1.	

In the conventional EGT, the new strategy F* would be added to the evolutionary game, leading to a symmetric 3 × 3 evolutionary game matrix, with the choice between F, MP and F* determined genetically. In this scenario, the residents who play F would not play differently against the F* mutant; hence the mutant would not be favoured against the resident. In contrast, the two-tiered approach allows male 1 to change its behaviour depending on whether it encounters a mutant male 2 or a resident one. In particular, suppose that male 1 is able to recognize that his payoff is higher when making peace than fighting and adjust its behaviour to take advantage of this higher payoff. Then the outcome of the behavioural game (1.2) would be the mutant male 2 playing F* and male 1 playing MP, which yields a payoff 3.5 > 1 to the mutant genotype. Hence, the mutant allele will be favoured against the resident, and natural selection will lead the population away from the prisoner’s dilemma behavioural game we started with.

The simple example above illustrates how the two-tiered conception of EGT opens up new ways to transform the evolutionary game. These possibilities bring with them a number of issues that need to be dealt with, such as how much information individuals have, and whether and how they can commit to various actions; we take up these issues as we introduce the general model below and also in §4. In the sections that follow, we introduce our formal two-tiered framework with goal-oriented dynamics at the behavioural tier and a population-genetic model at the evolutionary tier. We then present an analysis that generalizes the intuition from the motivating example above. In particular, we show under which conditions mutants can successfully invade a prisoner’s dilemma game by providing incentives for their opponents to cooperate. The resulting behavioural games are similar to the hawk–dove game. Furthermore, a genetic polymorphism can be maintained in the population, leading to a diversity of games and behavioural outcomes. We conclude with a discussion of our assumptions and the implications of our results.

2. THE FRAMEWORK FOR PAYOFF MATRIX EVOLUTION

Consider a social interaction between two individuals, such as the territorial competition described above. Individuals in such an interaction might have different roles (e.g. territory holder versus floater, or male versus female). The role of the individual might be determined genetically (e.g. male versus female) or environmentally (e.g. territory holder versus floater). Likewise, the actions available and the payoffs from those actions might be the same (resulting in a symmetric game, as in the game (1.1)) or different for the two roles. An allele that affects the entries in the payoff matrix can find itself in either role and in general might have different effects in each role.

To describe the frequency dynamics of alleles that affect the payoff matrix, we build a single-locus, diploid population genetics model. We have two alleles, A and B, and thus three genotypes: AA, AB and BB, which we index with \mathcal{A} , \mathcal{H} (for heterozygote) and \mathcal{B} , respectively. We denote a game by G_{ij} when genotype i is in role 1 and genotype j is in role 2. We assume that two individuals are randomly matched to play a game, and afterwards each is assigned a role in the interaction. We define ρ_{ij} as the probability that genotype i plays role 1 when paired with genotype j . By this definition, $\rho_{ij} = 1 - \rho_{ji}$, and thus $\rho_{ii} = 0.5$.¹

(a) The behavioural outcome

We assume for simplicity that the payoff from the focal game is the main determinant of an individual’s fitness. The outcome of the game is determined by a behavioural dynamics where individuals adjust how much time they allocate among the two actions as a function of their payoffs. In particular, denote the fraction of time the role 1 player allocates to its action 1 by x_1 ; similarly, denote by x_2 the fraction of time role 2 allocates to its action 1 ($x_1, x_2 \in [0,1]$). We will assume that the players adjust their allocations during behavioural time to maximize their own payoff (termed ‘individual play’ in [6]). Hence, the behavioural dynamic is given by

$$\frac{dx_1}{dt} = \frac{\partial u_1}{\partial x_1} \tag{2.1}$$

and

$$\frac{dx_2}{dt} = \frac{\partial u_2}{\partial x_2}, \tag{2.2}$$

where u_1 and u_2 are the payoffs to the role 1 and role 2 players; in a bimatrix game, the payoffs will be linear in x_1 and x_2 . We assume that the players will rapidly come to an equilibrium point of the behavioural dynamics. These dynamics are mathematically equivalent to a two-population bimatrix game, and therefore any stable equilibrium of the behavioural dynamics will correspond to a pure strategy Nash equilibrium (NE) of the game [14,15]. We label the outcome of game G_{ij} with Ω_{ij} , which is a vector with two components, representing the payoff to each player. Thus, $\Omega_{ij,1}$ is the payoff to the role 1 player, and $\Omega_{ij,2}$ is the payoff to role 2.

The behavioural dynamics allow individuals to discern their actual payoffs and adopt new actions when confronted with new games. The behavioural dynamic in our model is essentially a repeated game with a very fast succession of stage games, which allows individuals to adjust their actions. Note that the key assumption is that players can discern local (but not global) variation in their payoffs, and adjust behaviour ‘myopically’. As emphasized in §1, this behavioural plasticity is crucial to our model; in the absence of it, mutants that provide incentives to their opponents cannot ever benefit from doing that.

(b) The pairwise interaction model

We embed the behavioural dynamics in a diploid population genetic model that operates over micro-evolutionary time scale. In the electronic supplementary material, we provide the full recursion equations for the population genetic model, which are equivalent to

the ‘pairwise interaction model’ (PIM) of frequency-dependent selection [e.g. 16,17], with the interaction coefficients α_{ij} between genotypes i and j defined as

$$\alpha_{ij} = (\rho_{ij}\Omega_{ij,1} + (1 - \rho_{ij})\Omega_{ji,2}). \tag{2.3}$$

These interaction coefficients α_{ij} represent the expected payoff of genotype i from a pairing with genotype j , weighted by the probabilities of assuming either role. The recursion equations become:

$$\begin{aligned} \bar{w}q' = & f_A[f_A\alpha_{AA} + f_H\alpha_{AH} + f_B\alpha_{AB}] + 1/2f_H[f_A\alpha_{HA} \\ & + f_H\alpha_{HH} + f_B\alpha_{HB}], \end{aligned} \tag{2.4}$$

where \bar{w} is the mean fitness, q' the allele frequency in the next generation and f_i the Hardy–Weinberg frequency of genotype i when allele A is at frequency q (i.e. $f_A = q^2$, $f_H = 2q(1 - q)$, $f_B = (1 - q)^2$). The PIM always converges to an equilibrium and does not exhibit cyclic or chaotic dynamics when the interaction coefficients α_{ij} are non-negative [17], which is a reasonable assumption in our framework.

This full recursion equation simplifies under special circumstances, when $q \approx 0$ and $q \approx 1$, which give us the invasion and fixation conditions, respectively. Allele A can invade a population of BB homozygotes when $\alpha_{HB} > \alpha_{BB}$. Conversely, allele A can go to fixation when $\alpha_{AA} > \alpha_{HA}$. If no allele A satisfies the invasion condition, then allele B can be said to be evolutionarily stable (external stability *sensu* [18]).

Note that phenotypic ESS models are equivalent to a haploid, one-locus genetic model, which would have only two genotypes (A and B alleles). One can reduce our diploid model to a haploid one by adopting the convention of denoting one homozygote and the heterozygote genotypes by the alleles A and B, and considering the 2×2 evolutionary game matrix consisting of the interaction coefficients that correspond to these genotypes. Note that this modification of our model makes no difference for the invasion conditions, which can be interpreted in the context of standard evolutionary stability analysis.

3. STABILITY OF THE PRISONER’S DILEMMA

(a) The symmetric case

In this section, we apply our framework to a generic prisoner’s dilemma game, generalizing the example given in §1. We assume that the role distribution is symmetric (i.e. $\rho_{ij} = 0.5$ for all genotypes i and j). In addition, we assume the initially resident game is symmetric.

The game matrix between two homozygotes of the resident allele B, G_{BB} , is given by

		role 2		
		C	D	
$G_{BB} :$	role 1	C	r, r	s, t
	D	t, s	p, p	

(3.1)

where the actions C and D stand for ‘cooperate’ and ‘defect,’ and the following inequalities hold: $t > r > p > s$. The outcome of the game with individual play is $\Omega_{BB} = (p,p)$. Now, suppose a mutant allele A arises, which invests into changing the payoffs from the game. Specifically, assume that the mutant provides a ‘side-payment’ $\sigma > 0$ to its partner that it can only receive when it

plays C when the mutant is playing D, and pays a cost $\chi(\sigma) > 0$ for this side-payment. We assume that the cost reflects a decrease in a component of the mutant's fitness that is unrelated to the social interaction in question. For example, if the mutant allele reduces testosterone levels in the blood during territorial contests between two males, that might have a negative effect on the success of the mutant male in attracting a mate. By this assumption, the mutant pays the cost regardless of the eventual outcome of the interaction. Hence, mutant heterozygotes play the following games G_{HB} and G_{BH} with the resident homozygotes:

$$\begin{array}{l}
 G_{BH} : \quad \begin{array}{cc} & C & D \\ C & r, r - \chi(\sigma) & s + \sigma, t - \chi(\sigma) \\ D & t, s - \chi(\sigma) & p, p - \chi(\sigma) \end{array} \\
 \end{array} \tag{3.2}$$

$$\begin{array}{l}
 G_{HB} : \quad \begin{array}{cc} & C & D \\ C & r - \chi(\sigma), r & s - \chi(\sigma), t \\ D & t - \chi(\sigma), s + \sigma & p - \chi(\sigma), p. \end{array}
 \end{array}$$

(b) Invasion conditions

There are two possible cases: either $\sigma > p - s$ or $\sigma < p - s$. In the latter case, the NE is unchanged, and therefore such a mutant can never invade, since it is paying a cost and receiving no benefits. When $\sigma > p - s$, however, the outcomes of these two games are shifted relative to G_{BB} , and become $\Omega_{BH} = (s + \sigma, t - \chi(\sigma))$ and $\Omega_{HB} = (t - \chi(\sigma), s + \sigma)$. Thus, $\alpha_{HB} = t - \chi(\sigma)$, and the invasion condition $\alpha_{HB} > \alpha_{BB}$ becomes

$$t - p > \chi(\sigma). \tag{3.3}$$

In other words, for the invasion of a mutant making a side-payment, the side-payment has to be large enough to shift the NE, and the cost of this side-payment should be less than the benefit to be gained. With these criteria, the stability of a game depends critically on the relationship between the cost $\chi(\sigma)$ that the mutant pays to make a side-payment of σ . In the special case where the side-payment is zero-sum in nature (i.e. when $\chi(\sigma) = \sigma$), the invasion conditions are reduced to $t + s > 2p$. This means that whenever the total payoff from the new NE induced by the mutant allele is greater than the total payoff from the NE of the resident game, the resident game can be invaded by a mutant. On the other hand, if there is some 'inefficiency' in making side-payments (i.e. $\chi(\sigma) > \sigma$), invasion becomes more difficult, all else being equal. This situation can occur when the side-payment consists of a resource that increases one's payoff in an accelerating manner: if the individual with the greater resource makes the side-payment, its loss will be greater than the recipients' gain. On the other hand, gains can also occur, with $\sigma > \chi(\sigma)$, if the benefit from the resource exhibits diminishing returns to scale. This situation would facilitate the invasion of the resident game by the mutant A allele.

If a mutant can invade the game, what then is the consequence of an invasion (i.e. what does the game between mutant individuals look like)? The game between heterozygotes, G_{HH} is given below:

$$\begin{array}{l}
 G_{HH} : \quad \begin{array}{cc} & C & D \\ C & r - \chi(\sigma), r - \chi(\sigma) & s + \sigma - \chi(\sigma), t - \chi(\sigma) \\ D & t - \chi(\sigma), s + \sigma - \chi(\sigma) & p - \chi(\sigma), p - \chi(\sigma). \end{array}
 \end{array}$$

This payoff matrix looks very different from the prisoner's dilemma: both off-diagonal cells are now NE, but each player prefers a different one. This game is similar to the hawk-dove game [19]. Besides the presence of two alternative NE, the game G_{HH} also features reduced conflict of interest between the players, as shown in figure 1, which depicts the payoff polygons of the games G_{BB} given in equation (1.1) and G_{HH} with $\sigma = \chi(\sigma) = 1.5$. The payoff polygon is a plot of the different outcomes in the game in the payoff-space and the convex set resulting from linear combinations of these outcomes. The edges of the polygon running from the upper left to lower right-hand side denote the Pareto boundary: the set of outcomes upon which it is not possible to improve both players' payoffs simultaneously, also called efficient outcomes in economics. The length of this boundary can be taken as a measure of the potential conflict of interest: the longer the Pareto boundary, the greater the difference between the preferred outcomes of the two players. The invasion of the A allele shortens this boundary by moving the outcomes (C,D) and (D,C) closer together.² Furthermore, with the invasion of the mutant, the preferences of both individuals among the pairs (C,D) and (D,D), and (D,C) and (D,D) become concordant in game G_{HH} . This is another sense in which the allele A corresponds to reduced conflict between the players; see §4 for more on this issue.

Incidentally, a coordination game such as G_{HH} presents an additional question: which of the two NE will the behavioural dynamics result in? This question is beyond the scope of this paper and needs more detailed modelling of the behavioural interactions, including what initial actions individuals play. For the sake of simplicity, we assume that each NE is equally likely to be the outcome of behavioural dynamics. Thus, we take the expected outcome of the game to be the average of the two NE³, which is depicted by the larger circle on the Pareto boundary in figure 1b.

(c) Stable polymorphism

We now ask whether the mutant allele can also sweep to fixation. For simplicity, we assume that the effect of the mutant allele on the game is linear in the number of copies an individual carries (i.e. the mutant homozygote makes a side-payment of 2σ and incurs a cost of $\chi(2\sigma)$). (Our results are unchanged provided that the effect of the allele is monotonic in its copy number.) To calculate the interaction coefficient α_{HA} , we need the games between the heterozygote and mutant homozygote, G_{HA} and G_{AH} , which become

$$\begin{array}{l}
 G_{HA} : \quad \begin{array}{cc} & C & D \\ C & r - \chi(\sigma), r - \chi(2\sigma) & s - \chi(\sigma) + 2\sigma, t - \chi(2\sigma) \\ D & t - \chi(\sigma), s + \sigma - \chi(2\sigma) & p - \chi(\sigma), p - \chi(2\sigma) \end{array} \\
 G_{AH} : \quad \begin{array}{cc} & C & D \\ C & r - \chi(2\sigma), r - \chi(\sigma) & s + \chi(2\sigma), t - \chi(\sigma) \\ D & t - \chi(2\sigma), s + \sigma - \chi(2\sigma) & p - \chi(2\sigma), p - \chi(\sigma). \end{array} \\
 \end{array} \tag{3.4}$$

As in the game G_{HH} , both off-diagonal cells are NE in these games, such that the outcomes become $\Omega_{HA} = 1/2(s + t + 2\sigma - 2\chi(\sigma), s + t + \sigma - 2\chi(2\sigma))$ and $\Omega_{AH} = 1/2(s + t + \sigma - 2\chi(2\sigma), s + t + 2\sigma - 2\chi(\sigma))$. The

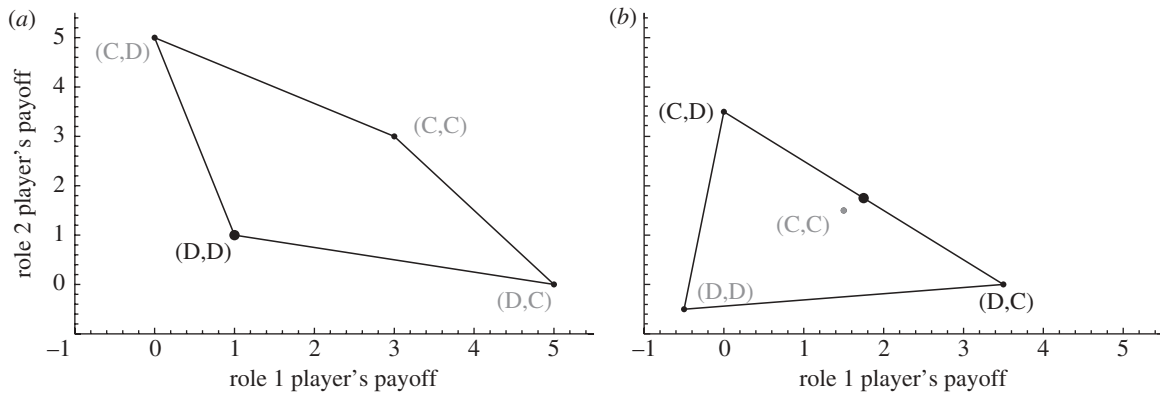


Figure 1. The payoff polygons for (a) the game G_{BB} (the resident prisoner’s dilemma behavioural game) and (b) the game G_{HH} (the game between the heterozygotes). Each corner of the polygons corresponds to a different cell in the payoff matrix, identified by the combination of actions. The points in grey are non-equilibrium outcomes, they are not played as the result of individual play dynamics. The outcomes in black are the result of individual play for each game. In (a), only the (D,D) outcome is the NE, whereas in (b), both (D,C) and (C,D) are possible outcomes. The large black circle in each panel denotes the expected payoff from the game to each role. In (b), this is the average of the two NE. In (a), the Pareto boundary is made up of the two line segments (D,C)–(C,C) and (C,C)–(C,D); in (b), the Pareto boundary is the line segment between the two NE. Thus, the expected payoff from the game G_{HH} lies on the Pareto boundary.

interaction coefficient is thus $\alpha_{HA} = 1/2(s + t + 2\sigma - 2\chi(\sigma))$. On the other hand, the game G_{AA} is

$$\begin{array}{l}
 G_{AA}: \\
 \begin{array}{cc}
 & C & D \\
 C & r - \chi(2\sigma), r - \chi(2\sigma) & s - \chi(2\sigma) + 2\sigma, t - \chi(2\sigma) \\
 D & t - \chi(2\sigma), s + 2\sigma - \chi(2\sigma) & p - \chi(2\sigma), p - \chi(2\sigma),
 \end{array}
 \end{array}
 \tag{3.5}$$

which yields an interaction coefficient $\alpha_{AA} = 1/2(s + t + 2\sigma - 2\chi(2\sigma))$. One can see that $\alpha_{HA} > \alpha_{AA}$ whenever $\chi(2\sigma) > \chi(\sigma)$ (i.e. when the cost is an increasing function of the side-payment). Hence, mutants that can invade the symmetric prisoner’s dilemma cannot also sweep to fixation.

The intuition behind this result is the following: since the homozygote games (G_{BB} and G_{AA}) are symmetric, and the effect of the mutant is role independent, a mutant that can shift the NE when in one role can automatically shift it in the other role as well. Thus, a heterozygote individual already is receiving all the benefits from making a side-payment, in return for a cost of $\chi(\sigma)$ only. The additional side-payment the homozygote makes does not further alter the NE and only results in the individual paying a higher, two-fold cost. Thus, heterozygotes enjoy an advantage against both homozygotes, resulting in a stable polymorphism.

What would such a polymorphic population look like? Figure 2 illustrates the stable polymorphism for a numerical case, where the payoff matrix G_{BB} is given by equation (1.1), the mutant A allele is characterized by $\sigma = \chi(\sigma) = 1.5$ and the polymorphic equilibrium is at $q \approx 0.39$. In this population, we would see all nine possible pairings of genotypes, corresponding to nine different games, and three different types of behavioural outcomes that lead to 13 unique NE payoff pairs overall. Because the genetic polymorphism is stable, the marginal fitnesses of the A and B alleles must be equal to each other. However, the same would not hold for the average fitness of individuals playing D and C at the behavioural equilibrium (who can be of various genotypes). For instance, in the example depicted in figure 2, the average fitness of individuals that play C at the behavioural equilibrium is approximately

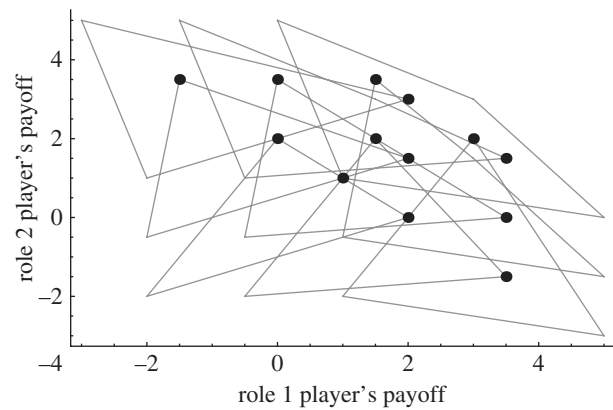


Figure 2. The effect of genetic polymorphism in the payoff matrices for the behavioural game. In this example, the resident behavioural game G_{BB} is given by the payoff matrix (1.1). Allele A is characterized by $\sigma = \chi = 1.5$; that is, it transfers a total of 1.5 payoff units (per copy of A) from the carrier of the allele to its social partner, when the former plays D and the latter C. At the polymorphic equilibrium of the population genetic model, there are three genotypes and hence nine different games being played in the population, which are represented by nine payoff polygons in this figure. The NE for each polygon is marked again with a large black circle. In total, there are 13 unique NE payoff pairs because the games between genotypes with at least one A allele have two NE (compare figures 1 and 3). The frequency of A alleles at the stable polymorphic equilibrium is $q \approx 0.39$.

1.01, whereas D-players’ average fitness is 2.62. Despite this marked difference between the average fitness of the two behaviours, both behaviours will persist in the stable polymorphism, since they are not determined by a simple genetic mechanism. However, a detailed genetic study on this population would nonetheless find a genetic component to which behaviour an individual converges towards, along with indirect genetic effects [20]. This population would therefore constitute a case where the alternative behavioural outcomes are

determined by both phenotypic plasticity and genetic polymorphism [21].

How robust is this polymorphism result to the assumption that the payoff matrix is determined by a single locus? To answer this question, we conducted simulations with multiple loci and recombination between those (see the electronic supplementary material). Our results indicate that with multiple loci polymorphisms are no longer inevitable, but substantial potential for polymorphic equilibria exist, provided the number of loci is not too great. Furthermore, the consequences for behavioural variation in the population hold true whenever the genetic polymorphism is maintained, regardless of the number of loci (see §4).

In the electronic supplementary material, we also discuss the cases of negative incentives and asymmetric games. For negative incentives, we can derive results that mirror the positive incentive case above. For asymmetric games, we find that there is a range of mutants that can both invade the population and proceed to fixation, in contrast to the symmetric case above (see the electronic supplementary material).

4. DISCUSSION

(a) Alignment of interests

We have presented a framework to model the evolution of the payoff matrix in the behavioural game, and applied this framework to study the evolution of payoff matrices starting from a prisoner's dilemma. Our first result shows that the prisoner's dilemma game can be invaded by mutants that provide incentives for cooperation. The evolutionary stability of a prisoners'-dilemma-type behavioural game depends on the nature of these incentives (e.g. how high a cost a mutant pays for a given change in payoffs). When individuals in one role are able to provide an incentive at a relatively low cost to themselves, it becomes easier for such mutants to invade, and for selection to align the interests of the two players. This result raises the question of whether prisoner's dilemma situations should be as ubiquitous as behavioural games in nature as they are in the theoretical literature. In particular, the prisoner's dilemma paradigm has attracted criticism for ignoring the social context of interactions [22], communicative mechanisms [23,24] and the possibility of direct benefits [25]. In one sense, our results are concordant with these objections, as we show that when animals are able to recognize their payoffs and have the potential to react to incentives, natural selection might lead away from the prisoner's dilemma and sustain a higher aggregate payoff through such incentives. On the other hand, we also show that this is not necessarily the case in all instances, and even when the population evolves away from the prisoner's dilemma by aligning the interests of the players, this alignment is not complete (see below).

The alignment of interests in our model occurs in two different senses. One is a 'local alignment': in the NE of the prisoner's dilemma behavioural game, both players play D, but each prefers the other to play C, while it keeps playing D itself. The invasion of the mutant allele shifts the NE (in game G_{HH}) by reversing the preference relations: at the new NE, one player plays D, the other C; the two players now concur in preferring that the second

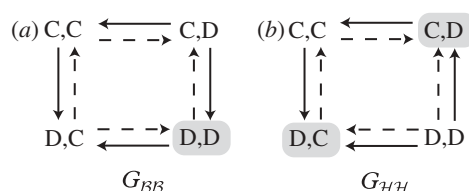


Figure 3. The movement diagram of the two games G_{BB} and G_{HH} . The arrows between each outcome (labelled by the players' actions) denotes whether each player prefers the outcome to another or not (solid arrows, role 1 player; dashed arrows, role 2 player). The NE for each game are marked by a grey rectangle. In G_{BB} (a), the leading and coming out of the NE are in opposite directions for the two players, signifying that they are in local conflict at this equilibrium. In G_{HH} , on the other hand, the arrows leading from (D,D) outcome to the two NE are in the same direction. For these movements, the players' interests are aligned. Note that the players are still in conflict about the movement between the NE and the (C,C) outcome.

player plays C instead of D. Thus, we can say that the interests of the players with regard to this move are aligned (figure 3). The second sense in which the interests are aligned can be seen in figure 1, where the difference between the best outcomes from each player's point of view is reduced by the invasion of the mutant allele. We call this 'global alignment', because it compares outcomes where both players have to engage in different actions. However, this alignment only occurs with positive incentives, and not with negative ones. The local alignment of interests is relevant for individual play (since coordinated change in actions is not possible under these dynamics), whereas global alignment is likely to be more relevant for when individuals can negotiate the outcome [6].

Closely related to our first result is the model by Worden & Levin [11], who also show that the population will evolve away from a prisoner's dilemma under such a scenario, which concurs with our findings. One important difference between the two models is that Worden and Levin assume that changing the payoff matrix is costless, which eventually leads to complete alignment of interests between the players, whereas interests are only partially aligned in our model.

The persistence of some payoff conflict between players suggests that behavioural mechanisms might evolve to resolve the remaining conflict, such as teamplay [6], other-regarding motivations [10,26] or positive response rules [3,4,7,27]. Conversely, the evolution of such a behavioural mechanism might mask the underlying payoff conflict, and hence might contribute to the evolutionary maintenance of it. Other types of behavioural decision rules, such as those aiming to maximize relative (instead of absolute) payoff, might also hinder the resolution of conflict at the payoff level.

Our model also relates to the theory of mechanism design (e.g. [28,29]), which deals with incentive schemes that induce self-regarding agents to reach desired outcomes, such as maximizing their employer's profits, or revealing truthful information in a bargaining situation. The local alignment of interests can be viewed as providing such an incentive scheme, especially since the average outcome of the mutant game is efficient (i.e. lies on the Pareto boundary; figure 1). The obvious difference

between our model and the mechanism design literature is that we do not assume a designer that has the power to change the game to produce results that fit its objectives. Instead, evolution acts as a 'blind mechanism designer', and the game emerges from the joint influences of the two individuals' genotypes.

A second difference is that mechanism design literature primarily deals with a problem that we did not include in our model, namely that individuals possess private information, and might have incentives to misrepresent it. The behavioural dynamics in our model rely on the individuals being able to observe and discern their actual payoffs, which include any side-payments that their partner's genotype induces (see below). This ability allows individuals to react optimally to new payoff structures brought on by mutant alleles. Without such behavioural plasticity, mutants that change the optimal course of action for their partners would have no hope of succeeding, since the only benefit such mutants enjoy results in changes in partner behaviour that they induce.

(b) *Polymorphism in games*

Our second result is that in symmetric interactions, a large class of alleles that affect the behavioural game will result in a protected polymorphism. Alleles in this class have monotonic effects on the phenotype (i.e. the behavioural game's payoff matrix) as a result of their copy number (e.g. homozygote mutants making twice the incentive payment and incurring twice the cost). If such a mutant can invade a symmetric interaction, heterozygote individuals will fare better against both homozygote genotypes, and hence a polymorphism will result. Taking the example given in the introduction of the paper, such a situation would manifest itself as a genetic polymorphism in genes regulating testosterone levels (or their receptors) in territorial males, but males of all types can be seen playing different actions, depending on the pairing of partners.

This finding suggests that diversity in behavioural games could be more common in nature than previously recognized, and might account for much of the diversity in behaviour that is observed. One potential example is the curious breeding system of the penduline tits (*Remiz pendulinus*) [30]. In this species, males build (with some help by the female) an elaborate nest, and either the male or the female, but not both, cares for the brood while the other deserts. However, about 30 to 40 per cent of all nests are deserted by both parents after the eggs are laid, with the consequence of the investment in nest-building and egg production being wasted. Since biparental care is missing, this pattern of uniparental care and mutual desertion cannot be explained by a simple genetic polymorphism in the caring tendencies. It can, however, be explained by a polymorphism in payoff matrices similar to the one arising in our model of the prisoner's dilemma game. The ancestral game of this clade is most probably more similar to the prisoner's dilemma, but with a negotiated behavioural outcome that results in biparental care. If for some reason the population evolves towards individual play, this would result in deserting being the behavioural outcome. From this point, a mutant such as that described in the symmetric prisoner's dilemma section can invade the population.

For the male's side, for example, the side-payments might be the effort spent by the male in producing a larger nest, which has indeed been found to increase the probability that the female cares for the brood, along with the brood size [31]. For the female, the mechanism of a possible side-payment might consist of laying larger eggs that grow faster and require less care. Hence, our model predicts that genetic polymorphisms for traits in each sex that affect costs and benefits from caring for the other sex will be present in the population and these polymorphisms will explain the behaviour of the two parents.

When the payoff matrix is determined with multiple loci, polymorphisms are not inevitable, since with multiple loci it becomes more probable that the minimum side-payment necessary to shift an NE will be a homozygote genotype (see the electronic supplementary material). Nonetheless, when the minimum side-payment requires a different number of A alleles on the two chromosomes, polymorphisms can be maintained with a small number of loci, and all the behavioural consequences of the genetic polymorphism continue to hold true in such cases. Thus, the question of how common polymorphisms in payoff structures are expected to be depends in part on how many loci are involved. For quantitative traits such as antler size, this number is likely to be high, so there is less potential for behavioural polymorphisms. On the other hand, regulatory genes that affect, for example, the expression levels of hormone or neurotransmitter receptors (e.g. [32]) or developmental pathways for morphology (e.g. [33]) can have a large effect on the phenotype. In those cases, polymorphic equilibria are more likely to be observed. Polymorphisms in the payoff matrix can also be maintained by other mechanisms, such as selection in heterogeneous environments (e.g. one that results in different G_{BB} matrices) coupled by gene flow. Regardless of the evolutionary mechanism of the maintenance of polymorphism, their effects on the behavioural diversity in the population will be the same.

(c) *Commitment and information*

Finally, our model raises questions about how individuals can commit to making side-payments that alter the behavioural game. In game theory, commitment problems arise when one party has more strategic flexibility than the other at some point in the interaction, and can use this flexibility to take advantage of the less flexible party. For instance, in the territorial interaction, a male might 'promise' a side-payment, but later withdraw this side-payment during game play. While such cases are in general possible (see below), there are two reasons why they do not happen in the current model setup. First, we assume the 'side-payment' to be a genetically determined trait, meaning that an individual cannot change its side-payment during the time scale of the social interaction. The side-payments can only change over evolutionary time through mutations and natural selection. This is a plausible assumption if the side-payment is a consequence of a phenotype such as the sequence of testosterone receptor gene; Obviously, the sequence of DNA is fixed during the time scale of a behavioural interaction, and hence the side-payment cannot be withdrawn.

On the other hand, even if the side-payments could be withdrawn, it would not pay for individuals to do that in our model. This is because we assumed that individuals are able to discern their actual payoffs and react quickly by adjusting their actions. Hence, even if a male withdraws the side-payment, its opponent would recognize this, and revert quickly to defection. Therefore, a side-payer that does not honour its promise would not receive the benefit of cooperation from its partner and would be at a disadvantage. In other words, side-payments in our model behave like simple contracts that are enforced through credible threats of retaliation.

The assumption that individuals can always react to each other and come quickly to an equilibrium is shared with previous models of two-tiered dynamics (e.g. [4,7,10]). This assumption simplifies the analysis greatly and allows clean, analytical results. Nonetheless, it is interesting to note what happens when equilibrium is not assumed. Doebeli & Knowlton [3] numerically evaluate the payoff to interacting partners during the transitory phase over a fixed number of rounds. Without any spatial structure, they find that the response rules eventually evolve to providing no benefits to the partner. Although Doebeli & Knowlton [3] do not address this issue directly, this reflects the inability of individuals to completely react to decreases in each other's investments due to the finite number of interaction rounds; hence mutants that take advantage of the cooperative types can invade. In the absence of effective retaliation, cooperative investments would unravel, similar to the way defection can be shown to be the only subgame perfect Nash equilibrium in a finitely repeated prisoner's dilemma [34]. Consistent with this conjecture, Doebeli & Knowlton [3] find that increasing the number of rounds increases the tendency for cooperative associations to spread in a spatially explicit model.

The assumptions of accurate information about payoffs and very quick reactions are likely to be not satisfied universally, but under some circumstances they will be reasonable approximations. For example, a male defending its territory is likely to be aware of how much territory it holds, and also have an estimate about what the expected breeding value of that territory is (our argument holds even if this estimate is imperfect). Then, we only require that the same male will also be aware of how much territory it loses during a territorial dispute while fighting versus not, and adjust the time spent on each action in order to maximize its expected payoffs. In general, our model is applicable to situations where individuals interact continuously in close proximity to each other.

On the other hand, commitment problems will become important when individuals have to make spatially or temporally separated decisions. Güth & Kliemt [35] investigated how internal mechanisms for commitment can evolve in a model of the 'trust game' where partners make sequential decisions. They show that if enough information is available about who is trustworthy and who is not, cooperative commitments can evolve because they provide incentives for their partners to cooperate, similar to the outcome in our model. Both types of mechanism are likely to play important and complementary roles in social evolution. The interplay between evolutionary dynamics of payoff structures and

behavioural mechanisms promises to be a fruitful avenue to understand the diversity of social interactions in nature.

We are grateful to J. Van Cleve, P. Iyer, L. Lehmann, Ç. Akçay, J. Fearon, B. Weingast, S. Gavrillets and V. Krivan for feedback on these ideas and the manuscript. We also thank three anonymous reviewers and Rob Boyd, who provided useful comments that clarified the argument and improved the manuscript. This research was supported in part by the Woods Institute for the Environment at Stanford University, and by a postdoctoral fellowship to E.A. from the National Institute for Mathematical and Biological Synthesis (an institute sponsored by the National Science Foundation, the US Department of Homeland Security and the US Department of Agriculture) through NSF Award no. EF-0832858, with additional support from The University of Tennessee, Knoxville.

ENDNOTES

¹An alternative specification would first specify the roles for the genotypes and then match individuals of different roles to play the game. That would lead to some changes in the equations, but the general methodology would be similar.

²The difference between the best and worst efficient outcomes for either player is $t - \chi(\sigma) - (s + \sigma - \chi(\sigma)) = t - s - \sigma$, whereas in the original game G_{BS} it was $t - s$.

³A mixed-strategy NE also exists, but it is unstable under individual play dynamics; therefore it will not be the outcome of the behavioural dynamics.

REFERENCES

- 1 Maynard Smith, J. 1982 *Evolution and the theory of games*. Cambridge, UK: Cambridge University Press.
- 2 Axelrod, R. & Hamilton, W. D. 1981 The evolution of cooperation. *Science* **211**, 1390–1396. (doi:10.1126/science.7466396)
- 3 Doebeli, M. & Knowlton, N. 1998 The evolution of interspecific mutualisms. *Proc. Natl Acad. Sci. USA* **95**, 8676–8680. (doi:10.1073/pnas.95.15.8676)
- 4 McNamara, J. M., Gasson, C. E. & Houston, A. I. 1999 Incorporating rules for responding into evolutionary games. *Nature* **401**, 368–371. (doi:10.1038/43869)
- 5 Taylor, P. D. & Day, T. 2004 Stability in negotiation games and the emergence of cooperation. *Proc. R. Soc. Lond. B* **271**, 669–674. (doi:10.1098/rspb.2003.2636)
- 6 Roughgarden, J., Oishi, M. & Akçay, E. 2006 Reproductive social behavior: cooperative games to replace sexual selection. *Science* **311**, 965–970. (doi:10.1126/science.1110105)
- 7 André, J. B. & Day, T. 2007 Perfect reciprocity is the only evolutionarily stable strategy in the continuous iterated prisoner's dilemma. *J. Theoret. Biol.* **247**, 11–22. (doi:10.1016/j.jtbi.2007.02.007)
- 8 Roughgarden, J. 2009 *The genial gene: deconstructing Darwinian selfishness*. Berkeley, CA: University of California Press.
- 9 Akçay, E. & Roughgarden, J. 2009 The perfect family: biparental care in animals. *PLoS ONE* **4**, e7345. (doi:10.1371/journal.pone.0007345)
- 10 Akçay, E., Van Cleve, J., Feldman, M. W. & Roughgarden, J. 2009 A theory for the evolution of other-regard integrating proximate and ultimate perspectives. *Proc. Natl Acad. Sci. USA* **106**, 19 061–19 066. (doi:10.1073/pnas.0904357106)
- 11 Worden, L. & Levin, S. A. 2007 Evolutionary escape from the prisoner's dilemma. *J. Theoret. Biol.* **245**, 411–422. (doi:10.1016/j.jtbi.2006.10.011)

- 12 Getty, T. 1987 Dear enemies and the prisoner's dilemma: why should territorial neighbors form defensive coalitions? *Am. Zool.* **27**, 327–336.
- 13 Akçay, Ç., Wood, W. E., Searcy, W. A., Templeton, C. N., Campbell, S. E. & Beecher, M. D. 2009 Good neighbour, bad neighbour: song sparrows retaliate against aggressive rivals. *Anim. Behav.* **78**, 97–102. (doi:10.1016/j.anbehav.2009.03.023)
- 14 Eshel, I. & Akin, E. 1983 Coevolutionary instability of mixed Nash solutions. *J. Math. Biol.* **18**, 123–133.
- 15 Hofbauer, J. & Sigmund, K. 1998 *Evolutionary games and population dynamics*. Cambridge, UK: Cambridge University Press.
- 16 Asmussen, M. A. & Basnayake, E. 1990 Frequency-dependent selection: the high potential for permanent genetic variation in the diallelic, pairwise interaction model. *Genetics* **125**, 215–230.
- 17 Schneider, K. A. 2008 Maximization principles for frequency-dependent selection I: the one-locus two-allele case. *Theoret. Popul. Biol.* **74**, 251–262. (doi:10.1016/j.tpb.2008.07.006)
- 18 Eshel, I. 1996 On the changing concept of evolutionary population stability as a reflection of a changing point of view in the quantitative theory of evolution. *J. Math. Biol.* **34**, 485–510. (doi:10.1007/BF02409747)
- 19 Maynard Smith, J. & Price, G. R. 1973 The logic of animal conflict. *Nature* **246**, 15–18. (doi:10.1038/246015a0)
- 20 Moore, A. J., Brodie III, E. D. & Wolf, J. D. 1997 Interacting phenotypes and the evolutionary process. I. Direct and indirect genetic effects of social interaction. *Evolution* **5**, 1352–1362. (doi:10.2307/2411187)
- 21 West-Eberhard, M. J. 2003 *Developmental plasticity and evolution*. Oxford, UK: Oxford University Press.
- 22 Noë, R. 1990 A veto game played by baboons: a challenge to the use of the prisoner's dilemma as a paradigm for reciprocity and cooperation. *Anim. Behav.* **39**, 78–90. (doi:10.1016/S0003-3472(05)80728-6)
- 23 Dugatkin, L. A., Mesterton-Gibbons and, M. & Houston, A. I. 1992 Beyond the prisoner's dilemma: toward models to discriminate among mechanisms of cooperation in nature. *Trends Ecol. Evol.* **7**, 202–205. (doi:10.1016/0169-5347(92)90074-L)
- 24 Noë, R. 2006 Cooperation experiments: coordination through communication versus acting apart together. *Anim. Behav.* **71**, 1–18. (doi:10.1016/j.anbehav.2005.03.037)
- 25 Hammerstein, P. & Leimar, O. 2006 Cooperating for direct fitness benefits. *J. Evol. Biol.* **19**, 1400–1402. (doi:10.1111/j.1420-9101.2006.01160.x)
- 26 Gintis, H., Bowles, S., Boyd, R. & Fehr, E. 2003 Explaining altruistic behavior in humans. *Evol. Hum. Behav.* **24**, 153–172. (doi:10.1016/S1090-5138(02)00157-5)
- 27 Wahl, L. M. & Nowak, M. A. 1999 The continuous prisoner's dilemma. I. Linear reactive strategies. *J. Theor. Biol.* **200**, 307–321. (doi:10.1006/jtbi.1999.0996)
- 28 Groves, T. 1973 Incentives in teams. *Econometrica* **41**, 617–631. (doi:10.2307/1914085)
- 29 Myerson, R. B. 1979 Incentive compatibility and the bargaining problem. *Econometrica* **47**, 61–73. (doi:10.2307/1912346)
- 30 Persson, O. & Öhrström, P. 1989 A new avian mating system: ambisexual polygamy in the penduline tit *Remiz pendulinus*. *Ornis Scandinavica* **20**, 105–111. (doi:10.2307/3676876)
- 31 Szentirmai, I., Komdeur, J. & Székely, T. 2005 What makes a nest-building male successful? Male behavior and female care in penduline tits. *Behav. Ecol.* **16**, 994–1000. (doi:10.1093/beheco/ari080)
- 32 Hammock, E. A. D. & Young, L. J. 2005 Microsatellite instability generates diversity in brain and sociobehavioral traits. *Science* **308**, 1630–1634. (doi:10.1126/science.1111427)
- 33 Chan, Y. F. *et al.* 2010 Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* **327**, 302–305. (doi:10.1126/science.1182213)
- 34 Osborne, M. J. & Rubinstein, A. 1994 *A course in game theory*. Cambridge, MA: The MIT Press.
- 35 Güth, W. & Kliemt, H. 2000 Evolutionarily stable co-operative commitments. *Theory Dec.* **49**, 197–221. (doi:10.1023/A:1026570914311)