Manipulated into giving: when parasitism drives apparent or incidental altruism

C. Patrick Doncaster,¹,* Adam Jackson,² Richard A. Watson²

¹Centre for Biological Sciences, Life Sciences Building 85, University of Southampton, Southampton SO17 1BJ, U.K. E-mail: cpd@soton.ac.uk

²Natural Systems Group, Electronics & Computer Science, University of Southampton, Southampton, SO17 1BJ, U.K. E-mails: aj3e10@soton.ac.uk; raw@ecs.soton.ac.uk

*Author and address for correspondence: C. Patrick Doncaster, Centre for Biological Sciences, Life Sciences Building 85, University of Southampton, Southampton SO17 1BJ, UK (cpd@soton.ac.uk)
Summary. Altruistic acts involve the actor donating fitness to beneficiaries at net cost to itself. In contrast, parasitic acts involve the actor extracting benefit from others at net cost to the donors. Both behaviours may have the same direct net-cost transferral of fitness from donor to beneficiary; the key difference between parasitism and altruism is thus who drives the interaction. Identifying the evolutionary driver is not always straightforward in practice, yet it is crucial in determining the conditions necessary to sustain such fitness exchange. Here we put classical ecological competition into a novel game-theoretic framework in order to distinguish altruism from parasitism. The distinction depends on the type of interaction that beneficiaries have amongst themselves. When this is not costly, net-cost transferrals of fitness from the donor are strongly altruistic, and sustained only by indirect benefits to the donor from assorative mixing. When the interaction amongst beneficiaries is costly, however, net-cost transferrals of fitness from the donor are sustainable without assorative mixing. The donor is then forced into apparent or incidental altruism driven by parasitism from the beneficiary. We consider various scenarios in which direct and indirect fitness consequences of strong altruism may have different evolutionary drivers.

Key words: biofilms, cooperative trading, density dependence, hawk-dove game, prisoner’s dilemma, snowdrift game

1. INTRODUCTION

An act of altruism confers a fitness advantage on others, which is strong altruism if the actor incurs a net fitness cost, and otherwise weak altruism [1,2]. The conferred advantage expresses a transferral of fitness from altruist to beneficiary, although the magnitude of the altruist’s loss may differ from the magnitude of the beneficiary’s gain. A parasitic act involves a costly transferral of fitness by the donor to a beneficiary, just as in strong altruism, with the crucial difference that the actor is the beneficiary and not the donor. The role of population
structure in sustaining these net-cost transferrals of fitness depends entirely on whether the individual driving the interaction is the donor or the beneficiary. Strongly altruistic acts of fitness transferral from donor to beneficiary can only be sustained by assorting mechanisms that cause the benefits of altruism to be visited disproportionately on other altruists or its costs to be cancelled by other indirect benefits [3-6]. In particular, assortative mixing that raises the probability of interactions with kin reduces the temptation to defect from self-sacrificing or other cooperative behaviours [1,7-9]. Parasitic acts of fitness transferral in contrast depend on the relative efficiencies of parasite-attack and host-defence mechanisms, with no inherent prerequisite for assorting mechanisms.

The initiator of an interaction, either as the donor in strong altruism or the beneficiary in parasitism, is readily identified in many cases. For example, worker castes in societies of eusocial insects clearly behave altruistically in renouncing their own reproductive potential for the benefit of the colony [10]. An ant clearly does not act altruistically in hosting a trematode fluke that will manipulate the ant’s behaviour to its own benefit [11]. The identity of the evolutionary driver is ambiguous in cases where the donor may gain indirectly from the interaction with a beneficiary, as in reciprocal or kin-selected altruism, while the beneficiary’s direct gain is set against a direct cost of conflict with other beneficiaries, as in competition amongst parasites for a limiting host resource. For example, a vampire bat may share its blood meal with others at net cost to its own fitness in acts of strong altruism, which can be sustained if the beneficiaries reciprocate or are kin [12]. Such indirect benefits do not exclude the possibility that parasitic manipulation by competing beneficiaries drives selection for the net-cost transferral of fitness, regardless of reciprocation or relatedness [13]. The behaviour is then what we will term either ‘apparent altruism’, if it brings no indirect benefits to the donor, or ‘incidental altruism’ if it brings indirect benefits which however have not motivated the interaction. Empirical studies of populations with structured interactions may have little
evidence with which to distinguish these alternatives, and evolutionary theory lacks a
common framework for comparing the social interaction of altruism to the exploitative
interaction of parasitism.

Whether the donor or the beneficiary drives the net-cost transferral of fitness has major
consequences for understanding the proximal causes of the evolution of altruism, because a
donor-driven transferral is not stable without assortative mixing, whereas a beneficiary-driven
transferral is stable even in a freely mixed population. The objectives of this paper are to
demonstrate that (i) a two-strategy game-theoretic analysis can distinguish strong altruism
from parasitism only by reference to the absolute values of payoffs; (ii) the beneficiary may
drive selection for the net-cost transferral of fitness in parasitism, even if the donor drives
selection for the assortative mixing that sustains strong altruism.

We construct a formal game theoretic framework to set the evolution of cooperation in a
broader ecological context capable of distinguishing altruism from parasitism. Reformulation
of standard models from ecology within game theory shows that costly interactions amongst
beneficiaries make their interactions with the donor exploitative, and consequently sustainable
with or without assortative mixing of players. A net-cost transferral of fitness from donor to
beneficiary may be driven by parasitic manipulation, regardless of whether the donor achieves
indirect fitness gains from assortative mixing. This is an important clarification to current
understanding of strong altruism, which assumes that the evolution of its net-cost transferrals
of fitness is driven by the indirect fitness benefits to the donor resulting from assortative
mixing [1,3-7]. Competition amongst kin is then predicted to reduce selection for altruistic
behaviour [8,14-17]. Our game-theoretic framework for parasitism reveals how mutual
competition between parasitic beneficiaries can create selection for altruistic donors.
2. GAME-THEORETIC FRAMEWORK FOR TWO STRATEGIES

The problem of how traits for strong altruism can be favoured by natural selection is embodied in a two-strategy game between two players. Each player is allowed to choose between a strategy of social cooperation and a strategy of selfish defection. It then receives a payoff that depends on its own and the other’s strategy, as shown in table 1. Mutual cooperation is not a stable strategy when its $R$ (‘Reward’) payoff to each Cooperator is less than the $T$ (‘Temptation’) payoff to a Defector for unilateral defection. Moreover, unilateral cooperation is not a winning strategy when its $S$ (‘Sucker’) payoff to the Cooperator is less than the $P$ (‘Penalty’) payoff to each Defector for mutual defection. This is the Prisoner’s Dilemma (PD). Players of the PD find that unilateral defection against a Cooperator pays better than mutual cooperation, and mutual defection pays better than unilateral cooperation.

The game is a social dilemma if the payoffs create a tension between individual welfare and collective welfare given by the payoff sum for both players [18]. It is a useful analogy for social evolution theory when the same-strategy payoffs $R$ and $P$ take the form of fitness increments, and cross-strategy payoffs $T$ and $S$ can involve either a decrement or an increment to fitness. For example, given a cooperative behaviour that confers fitness benefit $b$ at net cost $c$, a game may have payoffs $T = b$, $S = -c$, $R = b - c$, $P = 0$ [19,20]. The PD then represents the problem that strong altruism, with $T > 0 > S$, is not a stable outcome amongst freely interacting players. Any number of other relationships are possible between the fitness payoffs $T$, $S$ (or $b$, $c$) and $R$, $P$, depending on the scenario [20-26].

In order to bring exploitative interactions into game theory for two strategies, we distinguish three different categories of games, according to whether the same-strategy payoffs $R$ and $P$ are (i) both non-negative, or (ii) $R$ is positive and $P$ is negative, or (iii) both are non-positive. All previous applications of two-strategy games to evolution have assumed a non-negative Reward payoff. Whereas social evolution theory conventionally takes a non-
negative Penalty payoff [19-22], evolutionarily stable conflict can take positive Reward and 
negative Penalty (e.g., in Hawk-Dove or Hawk-Mouse variants on Defector-Cooperator 
games [23,24]. A zero Reward and negative Penalty can apply to games of selection for the 
same trait (‘kind selection’) encompassing antigreenbeard traits [25,26]. Here we show that a 
negative Reward and Penalty are appropriate to games involving competitive exploitation 
amongst individuals. Our analysis of games across all three categories yields the complete set 
of net-cost fitness transferrals from donor to beneficiary in altruistic and/or parasitic (or 
predatory) interactions, regardless of how the magnitude of the donor’s loss converts to the 
beneficiary’s gain.

We measure payoffs in the conventional way, as the change in fitness to the actor 
resulting from the interaction with the recipient [2]. Payoffs are thereby calibrated against the 
unavailable option of no interaction, corresponding to zero change [24]. We apply this 
calibration to symmetric games, in which only four payoffs are possible from pairwise 
interactions amongst two types of player. We enumerate all outcomes for these games at three 
scales: (i) two strategies for two players (‘strategic players’), (ii) two phenotypes in a single 
population (‘replicator dynamics’), and (iii) two genetically isolated populations (‘ecological 
dynamics’).

3. TWO-PLAYER DYNAMICS OF SOCIAL AND EXPLOITATIVE GAMES

The Prisoner’s Dilemma is one of four canonical games between a Cooperator and a Defector 
that describe the complete 4-dimensional parameter space of payoffs. That is, each of these 
games is defined by the relative values of $T, R, S$ and $P$ in the table-1 payoff matrix, from 
which game theory determines the winning strategy (or equilibrium strategies). Figure 1(a) 
shows thresholds of the payoff matrix for the four canonical games of Prisoner’s Dilemma 
(PD), Snow-Drift (SD, including evolutionarily stable Hawk-Dove or Hawk-Mouse games), 
Harmony Game (HG), and Stag Hunt (SH). The coloured domains in figure 1(b) show the
game outcomes. The Prisoner’s Dilemma is won by the Defector to the exclusion of the Cooperator; the Snow-Drift game sustains both Defector and Cooperator (which may take the roles respectively of Hawk and Dove/Mouse); the Harmony Game is won by the Cooperator to the exclusion of the Defector; the Stag Hunt game has bi-stability [27].

Comparisons between games are facilitated by reducing the dimensionality, for example by normalizing the rankings against a 1-unit Reward and zero Penalty [28,29]. In order to facilitate the mapping of social with exploitation games, we introduce a more flexible normalization that fixes the Reward at a constant $R$ units and the Penalty at a lower value of $P = \pm 1/R$. The condition $R > P$ encompasses three categories of games, illustrated in figures 2(a)-(c). These are: (a) a positive Penalty that is less beneficial than a positive Reward (shown with $R > 1$, $P = 1/R$); (b) a negative Penalty and positive Reward ($P = -1/R$); (c) a Penalty that is more costly than a negative Reward ($0 > R > -1$, $P = 1/R$). These alternatives preserve the integrity of payoff rankings that define each game, such that the winning strategy remains unchanged by the sign of the Reward and Penalty (e.g., cooperation never wins in the Prisoner’s Dilemma). For all three categories, their four games have unconfined domains, extending indefinitely beyond the axes. Social dilemmas arise only below the angled pink lines in figure 2, where collective welfare pays better than individual welfare ($2R > T + S$), and the benefits of cooperation conflict with those of defection (greed prevails with $T > R$ or fear prevails with $P > S$), always assuming $R > P$ [18]. The social dilemma ignores any specifics about who initiates interactions, or their mechanisms or functions.

Assigning a positive value to $R$ puts the origin of the Temptation-Sucker plane within the SH domain when $P \geq 0$ (figure 2(a) for $P = 1/R$). The quartiles surrounding the origin describe alternative payoffs for the Cooperator in terms of the four modes of social behaviour in the Hamiltonian classification [2,30,31]: +/- mutual benefit, +/- strong altruism (stippled), –/- spite, or –/+ selfishness. The Cooperator in the mutual-benefit quartile is weakly altruistic.
across the segment $T > S > 0$, defined by its altruistic transferral of fitness to the Defector ($T > S$) combined with its net fitness gain ($S > 0$) distinguishing it from strong altruism. In contrast, the Cooperator in the strong-altruism quartile is strongly altruistic throughout, defined by $T > 0 > S$, with the negative $S$ expressing its net fitness loss relative to non-interaction. The Cooperator strategy is not a stable outcome in the strong-altruism quartile. Selfishness is stable in the Defector-only domain of PD, and mutual benefit is stable throughout the coexistence domain of SD.

Given the possibility of games with a negative $P$, the origin of the Temptation-Sucker plane is not constrained to the SH domain. In fact, the only domain from which it is excluded is PD, because that would require $P > 0 > R$. Figure 2(b) shows the origin in HG for games having same-strategy interactions that are beneficial for the Cooperator but costly for the Defector ($R > 0 > P$). The immediate consequence of moving the origin out of SH is that the +/- interaction is no longer confined to the Defector-only outcome of PD. It becomes an alternative to the +/- interaction as a coexistence outcome within SD (stippled green sector), and is sustained in the Cooperator-only HG (stippled blue sector).

Figure 2(c) shows the origin in the SD domain for games between two strategies that both have negative same-strategy payoffs. The Cooperator strategy in this scenario may be cooperative only in the relative sense of its same-strategy interaction (Cooperator-Cooperator) being less costly than the Defector’s same-strategy interaction (Defector-Defector). Now all four types of interaction behaviour have coexistence outcomes in SD.

A Defector strategy with negative $P$ gives coexistence outcomes the character of Hawk-Dove or Hawk-Mouse conflicts [24]. The +/- interaction in the SD domains of figures 2(b)- (c) more reasonably describe a manipulator-victim or parasite-host relationship than a selfish-altruist relationship of the sort that applies to PD in figure 2(a), even for the social Cooperator in figure 2(b).
For any social dilemma, and regardless of any normalization of same-strategy payoffs, net-cost transferrals of fitness are excluded from SD only by \( P \geq 0 \). A net cost transferral of fitness therefore has alternative domains set by \( P \). It arises in PD amongst players with non-negative \( P \), where it is interpreted as strong altruism and cannot be sustained without assortative mixing (e.g., kin selection). In the particular case of \( P = 0 \), and only in this case, unilateral cooperation in PD is synonymous with strong altruism. Alternatively, the net-cost transferral of fitness arises in SD amongst players that have negative \( P \), where it is sustained in freely mixing populations as victimization by an aggressor. Game theoretic treatments of social evolution conventionally equate the cross-strategy interaction in PD with altruism [5,19,32-36]. In order to demonstrate fully the danger of assuming this equivalence, we further develop the analysis of games with negative \( P \). The following two sections explain the net-cost transferrals of fitness in SD in terms of replicator dynamics and ecological dynamics.

### 4. WITHIN-POPULATION DYNAMICS OF SOCIAL AND EXPLOITATIVE GAMES

Two-strategy games for two players are extended to two-strategy games for a population of players for the purposes of modelling evolutionary change by frequency-dependent selection. Here we use standard replicator dynamics to describe evolutionary games for a population of two phenotypes, in which their relative frequencies in the population determine individual fitness [37-39]. The equivalent ecological model then reveals the role of exploitative interactions in sustaining net-cost transferrals of fitness between phenotypes.

Let strategies \( E_1 \) and \( E_2 \) have frequency-dependent expected payoffs \( (A \cdot x)_1 \) and \( (A \cdot x)_2 \) for playing Cooperator and Defector respectively, where \( A \) is the table-1 matrix of fitness payoffs, and \( x \) is the vector of fractions \( x_1 \), \( 1 - x_1 \) describing the relative frequencies of Cooperator and Defector. The evolutionary success of strategy \( E_i \) is expressed as the difference between its fitness and the average fitness: \( (A \cdot x)_i - x \cdot A \cdot x \) [38]. For a large
population susceptible to continuously changing frequencies, this difference determines the per capita rate of change \( \dot{x}_i / x_i \) over time \( t \). The replicator equation for each of two strategies is then

\[
\dot{x}_i = x_i \left[ (A \cdot x) - x \cdot A \cdot x \right].
\]  

(4.1)

Solving for \( x_i \) in equation (4.1) at \( \dot{x}_i = 0 \) gives the evolutionarily stable frequency \( x_i^* \) of strategy \( E_i \). This is 0 or 1 if \( P \geq S \) and/or \( R \geq T \), and otherwise takes positive stable fractions \( x_1^* = (S - P)/(S - P + T - R) \) and \( x_2^* = (T - R)/(S - P + T - R) \) [38]. These relative frequencies equal the probabilities of playing each strategy in the two-player game [24].

The two-phenotype replicator equation (4.1) is numerically equivalent to a single-genotype ecological model with the Lotka-Volterra form

\[
\dot{y} = y \left[ (S - P) + (R - T) \cdot y \right],
\]  

(4.2)

where \( y = x_1/x_2 \) and \( x_2 \cdot t \) defines the time-scale [38]. A Cooperator strategy can invade the genotype on condition \( S > P \), whereupon its ratio with a Defector strategy grows logistically on condition \( T > R \), to equilibrium \( y^* = (S - P)/(T - R) \).

Figures 2(a)-(c) model the categories of equilibrium outcome that apply equally to equations (4.1) and (4.2) on the \( T \cdot S \) plane for positive and negative \( R \) and \( P \), with \( P = \pm 1/R \).

Just as for two players, figures 2(b)-(c) show that a homogeneous population can sustain net-cost transferrals of fitness from Cooperator to Defector phenotype provided \( P < 0 \). The stippled green sector of the \( T \cdot S \) plane sustains this type of interaction, and in a social dilemma below the angled pink line. The transferral of fitness from Cooperator to Defector is driven by the Defector sustainably exploiting the Cooperator in homogeneously mixed interactions, so without requirement for indirect fitness benefits to the Cooperator as in strong altruism.

An example illustrates the danger of assigning \( P \geq 0 \) to a scenario that may have costly same-strategy interactions. Meat sharing amongst non-kin chimpanzees has potential
explanations in reciprocal altruism or manipulative begging [40]. Consider a population
containing a Cooperator phenotype that donates meat to others at net cost to its own fitness,
and a Defector phenotype that benefits from the donated meat without reciprocating. The
meal-sharing process may be one of strong altruism by Cooperators or parasitic manipulation
by Defectors. Strong altruism is only sustainable if population structure or other mechanisms
of assortative mixing direct the benefits of altruism disproportionately to other altruists (or to
kin). In contrast, manipulation by Defectors has no structural pre-requisite given a negative
payoff for their same-strategy interactions ($P < 0$). Competition for mutually limiting
resources presents a natural interpretation for $P < 0$, when Defectors depress each other’s
reproductive success below the intrinsic success achievable in the absence of others. Thus in
equation (4.2), $P < 0$ allows incursion into the Defector population by Cooperators that
themselves suffer competition with the Defectors ($S < 0$), always provided $S > P$. Under these
conditions, manipulative behaviour in an unstructured population once established may
subsequently favour selection on food donors to suffer it only from other donors (or kin). The
Cooperator phenotype then favours population structures that facilitate reciprocal (or kin-
selected) donation of fitness. The interaction is nevertheless incidental altruism driven by
parasitism for as long as $P < 0$ sustains the net-cost transferral of fitness even without such
population structures. Strong altruism cannot be the evolutionary driver of the +/- interaction
itself, it can only drive selection for the assortative mixing that confers indirect benefits on
donors through reciprocation (or kin beneficiaries). In effect, donors do not give benefit to
others in order to receive it back reciprocally or via kin, the benefit is taken from them by
others. To the extent that they are able, however, it makes sense for donors to limit the theft to
reciprocators or kin, for example by limiting dispersal (a form of ‘social niche construction’
[41,42]). Similarly, intraspecific brood parasitism in birds can function as helping when the
host is a relative [43]. Any net fitness cost to an individual in hosting a relative’s egg is
incidental altruism if the behaviour is sustained with no lower threshold for the indirect benefit to the host through relatedness.

5. BETWEEN-POPULATION DYNAMICS OF EXPLOITATIVE GAMES

The 1-dimensional system of equation (4.2) for a single population readily expands to accommodate two genetically isolated populations competing within and between their genotypes to exploit limiting resources. Standard Lotka-Volterra coupled rate equations describe the 2-dimensional population dynamics for a closed community:

\[
\frac{\dot{n}_i}{n_i} = r_i \cdot n_i \left[ 1 + \frac{(\alpha_{ij} \cdot n_j - n_i)}{k_i} \right].
\]

For each Genotype-\( i \), rate \( \dot{n}_i \) describes continuous change over time in abundance \( n_i \), where \( k_i \) is its carrying capacity in the absence of the other genotype. Population change is determined by the intrinsic rate of increase per capita \( r_i \), moderated by the density of both populations, with interaction coefficient \( \alpha_{ij} \) measuring the impact on Genotype-\( i \) from Genotype-\( j \) relative to the normalized impact \( \alpha_{ii} = -1 \) on Genotype-\( i \) from its own type. For example, Genotype-\( i \) takes \( \alpha_{ij} < 0 \) if its population growth is impeded by competition from Genotype-\( j \). The negative payoff to Genotype-\( i \) from the interaction then counts as a cost relative to intrinsic fitness \( r_i \) before any interactions.

In terms of a game, payoffs are measured relative to fitness \( r_i \) available prior to playing the game. The calibration of payoffs is manifested explicitly by viewing a Lotka-Volterra model for \( G \) genotypes as an equivalent replicator model for \( G + 1 \) phenotypes [38]. Accordingly, the equation-(5.1) model for genotypes \( i = 1, 2 \) translates to replicator equation

\[
\dot{x}_i = x_i \left[ (A \cdot x)_i - x \cdot A \cdot x \right]
\]

for phenotypes \( i = 1 \) to 3 with \( x_1 + x_2 + x_3 = 1 \), taking payoff matrix \( A \):
The bottom row of $A$ takes zeros, reflecting the absence of a third active strategy. $A_{13}$ and $A_{23}$ are then payoffs to Cooperator and Defector respectively without interactions.

Solving for $n_i$ in equation (5.1) at $\dot{n}_i = 0$ gives its stable equilibrium population size:

$$n_i^* = k_i + \alpha_{ij} \cdot n_j,$$

(5.3)

Given $n_j^* = k_j$ when $n_i^* = 0$, solving for $\alpha_{ij}$ in equation (5.3) yields the necessary and sufficient condition for positive equilibrium abundance of Genotype-$i$ (i.e., $n_i^* > 0$):

$$\alpha_{ij} > -k_i/k_j.$$  

(5.4)

When both populations meet condition (5.4), equation (5.3) yields stable $n_i$ at system equilibrium:

$$n_i^* = \left( k_i + \alpha_{ij} \cdot k_j \right) / \left( 1 - \alpha_{ij} \cdot \alpha_{ji} \right).$$  

(5.5)

These predictions describe the equilibrium outcomes of ecological processes defined by equation (5.1) for a two-strategy game played by two populations. They have been thoroughly analysed in ecological theory [44-46], though not at all in game theory.

Figure 2(d) shows how the $\alpha$ plane of interaction coefficients partitions into domains of equilibrium persistence for one or both genotypes on condition (5.4). The green domain encompasses stable equilibrium coexistence, with the populations growing monotonically, or in damped oscillations, to positive $n_1^*$ and $n_2^*$. The white domain encompasses bi-stability. All four outcomes on the $\alpha$ plane have exactly corresponding outcomes on the Temptation-Sucker plane of figure 2(c), with assignment of cross-strategy payoffs $S = \alpha_{12}$ and $T = \alpha_{21}$. The mutual payoffs $R = -k_2/k_1$ and $P = -k_1/k_2$ enumerate the greater efficiency of G1 relative to G2 in costly resource exploitation. They are forcibly non-positive because the carrying capacities $k_1$ and $k_2$ take non-negative values. With the Cooperator genotype G1 having the greater
carrying capacity, figure 2(d) has \(-k_2/k_1 > -k_1/k_2\) just as figure 2(c) has \(R > 1/R\) for \(0 > R > -1\) and \(P = 1/R\).

The \(\alpha\)-plane origin must lie within the coexistence domain, given negative \(R\). Figure 2(d) shows its surrounding quartiles, which determine the identities of freely interacting genotypes. The four alternatives are \(+/+\) mutualistic, \(+/-\) parasite-host, \(-/-\) competitors, and \(-/+\) host-parasite [47]. Given the costly nature of same-genotype interactions, all cross-genotype interactions involve forms of selfish exploitation, whether to mutual benefit or cost, or to the benefit of one at the cost of the other. Within the domain of coexistence, positive or negative \(a_{ij}\) for Genotype-\(i\) are expressed in larger or smaller \(n_i^*\) relative to \(k_i\), as determined by equation (5.5). Thus a genotype has parasitic status if it achieves higher equilibrium abundance in the presence of the other genotype than on its own, resulting from a transferral of fitness at net-cost to the other, which therefore has host status. For example, the dynamic of cuckoos parasitizing the nests of other bird species is sustained in the model by the surrogate parent genotype G1 having a sufficiently higher carrying capacity \((k_1 > k_2)\) to compensate for its net cost in fitness \((a_{12} < 0)\) from provisioning nestlings to the net benefit of the cuckoo chick \((a_{21} > 0)\) with genotype G2. As a two-strategy game, the net-cost transferral of fitness from the surrogate parent to the net benefit of the cuckoo chick \((T > 0 > S)\) is sustained by virtue of a sufficiently negative \(P\) to accommodate \(0 > S > P\) in the SD domain. Whilst cuckoos are unmistakably parasitic, net-cost transferrals of fitness between symbiotic species can be sustained as altruism when indirect benefits accrue to relatives of the donor [48]. In such cases, a game-theoretic analysis of the sort illustrated in figure 2(d) can identify apparent or incidental altruism in which it is parasitism that drives the direct fitness benefits.

In game-theoretic terms, the population dynamics of two resource-limited genotypes can only encompass all four games when the origin of the \(T-S\) plane is in the coexistence domain of SD (figures 2(c)-(d)). To locate the origin in SH or HG (as figures 2(a)-(b)) would mean...
that inter-genotype interactions are necessary to the persistence of one or both genotypes, which in turn would rule out the existence of SH and HG and/or PD games. For example, a predator G2 that depends on a prey G1 has standard Lotka-Volterra dynamics:

$$\dot{n}_1 = c_1 \left[1 - n_1 + \alpha_{12} \cdot n_2 \right] \cdot n_1 - d_1 \cdot n_1, \quad \dot{n}_2 = c_2 \cdot \alpha_{21} \cdot n_1 \cdot n_2 - d_2 \cdot n_2.$$  \hspace{1cm} (5.6)

The impacts of predation are measured in a fitness cost $\alpha_{12} < 0$ to each individual of the prey population $n_1$, and a corresponding fitness benefit $\alpha_{21} > 0$ to each individual of the predator population $n_2$. The $\alpha$-plane origin lies in the G1-only domain (equivalent to the T-S plane having its origin in HG), as a result of predator persistence requiring that $\alpha_{21}$ exceeds a positive threshold $R$ (specifically, $n^* = 0$ at $R = d_2/(c_2 \cdot k_1)$ where $k_1 = 1 - d_1/c_1$). Since G2 cannot persist without G1, the $\alpha$ plane has no domains equivalent to the game-theoretic PD and SH (in effect, $P \rightarrow -\infty$).

### 6. WHEN IS A NET-COST FITNESS TRANSFERRAL NOT STRONG ALTRUISM?

The game-theoretic homology between victim or host responses to manipulation and strongly altruistic acts is not a trivial issue of re-labelling an altruist as victim. It speaks directly to questions about the evolutionary origins of cooperation, in ways that social game theory cannot when it confines itself to games with $P \geq 0$, and therefore to already cooperative players. For example, a strain of bacterium G1 that manufactures a public good in the form of an extra-cellular polymeric biofilm may suffer its exploitation by a strain G2 that contributes nothing to the public good [49]. These types of interactions are commonly construed as forms of strong altruism requiring explanation in terms of population structure [25,50-51]. However, equations (5.1)-(5.5) predict stable coexistence of the two strains as +/- parasite-host in freely-mixing populations, provided G1 sufficiently compensates for its competitive inferiority with superior carrying capacity. A detailed physiological model of the growth in thickness of biofilms has predicted a potential competitive advantage to biofilm production in...
sustaining access to oxygen, analogous to vertical growth in plants towards light [52]. Surface layers of biofilm suffocate deeper layers, in interactions between strains consistent with sustainable $+/-$ competition or $+/-$ parasitism in SD (figures 2(c)-(d)). Other acts of apparent altruism may also have testable alternative explanations in terms of exploitation. For example, plants that evolve to refrain from overshadowing their neighbours can be altruists [53]. An alternative driver in competitive exclusion should not be ruled out, however, unless the same-strategy interaction amongst beneficiaries has a non-negative fitness payoff.

Table 2 enumerates the full set of seven different social dilemmas and one HG that have cross-strategy interactions with a net-cost payoff to Cooperators and net-benefit payoff to Defectors. These include all combinations of game (PD, SD, HG, SH) with positive or negative $R$ and $P$ and $+/-$ interactions between strategies, as stippled in figure 2. The three occurrences of outcomes in PD, and the two in SD, emphasize that the cross-strategy interactions in PD and SD are not synonymous with strong altruism and mutual benefit respectively, except in the restrictive case of $P = 0$. The many alternative outcomes of $+/-$ interactions show that to assume non-negative same-strategy interactions ($R > P \geq 0$), if inappropriate for the scenario at hand, will result in a host response to manipulation or parasitism in SD being mistaken for strong altruism in PD.

Distinguishing between beneficial and costly same-strategy interactions requires measuring the fitness increment or decrement due to $R$ or $P$ above or below the intrinsic fitness available without interaction. For some eusocial insects or vertebrate societies in which lone individuals have low or zero survival, same-strategy interactions clearly bring benefits. In other cases, and particularly amongst microorganisms, it is not trivial to know whether same-strategy interactions are beneficial or costly relative to no interaction, because the reproductive success measurable in natural environments is usually that achieved in the presence of intra- and inter-specific interactions. To assume a non-negative $P$, however, can
lead to multiple misconceptions. These include wrongly identifying or predicting altruism, or wrongly interpreting the evolutionary driver of the net-cost transferral of fitness as kin- selected or reciprocal altruism.

7. DISCUSSION

The mapping of replicator and Lotka-Volterra dynamics onto a common plane of payoffs for cross-strategy interactions has defined the conceptual difference between altruistic and parasitic interactions. We have shown how negative $R$ and/or $P$ are naturally conceived in ecological scenarios of a two-phenotype population and of independently self-replicating populations. They provide evolutionary game theory with an ecological framework for setting net-cost transferrals of fitness within the SD domain of stable coexistence. Moreover, they open ecological theory to game-theoretic analysis.

Conventional game theory for social evolution assigns non-negative payoffs to same-strategy interactions, which are then consistent with the convention to consider social behaviours as any interactions amongst same or cross strategies that have fitness consequences for actor and recipient [2]. In order to include competitive interactions within two-strategy games, we find it useful to consider a strategy as ‘social’ only if its same-strategy interaction has a non-negative payoff, making it at least as good as no interaction. The alternative to a social strategy is then an ‘unsocial’ strategy defined by a costly same-strategy interaction, making it an adverse encounter typical of mutual competition for limiting resource. Two-strategy games for two or more players assume that the game gets played, whether the players have elected or been forced to play it. Given the distinction between social and unsocial strategies, it makes sense to interpret a game as elected if both strategies are social. Then neither strategy incurs a cost by playing its own type, even if one or both may be susceptible to invasion by the other. Alternatively, a game is enforced if at least mutual defection has a negative payoff. Then at least the strategy of mutual defection incurs a cost by
playing, as for example the impact of same-strategy competition on the intrinsic fitness prior to density effects. Enforcement is thus by an extrinsic process, such as competition for limiting resource.

It has long been recognized that competition within population structures can reduce or cancel the net selective advantage of altruistic behaviour [8,14-17]. For kin structures, this balancing of effects depends on competition being coupled to relatedness [8, 15,54,55]. Our analysis is consistent with these results insofar as the strong altruism enacted by a Cooperator in a PD game requires a larger sacrifice in the presence of the costly same-strategy interactions typical of density-regulated competition (i.e., a more negative $S$ in the PD domains of figures 2(b)-(d) than in figure 2(a)). The PD prediction of a Defector-only outcome applies to the case of homogenous competition and relatedness, and it is only the decoupling of relatedness from competition that can allow the Cooperator to direct altruism at kin. Crucially, however, the game theoretic representation of competition has shown how a victim of parasitic manipulation also functions as an incidental altruist if it allows itself to be victimized only by kin, in which case competition may motivate altruism via parasitic manipulation. This is a distinct mechanism from the policing of altruism by coercion, which is driven by altruists to enforce altruistic behaviours [4,56].

In the literature on the evolution of cooperation, the conventional application of game theory is to discrete phenotypes with replicator dynamics in elected games. Ecological contexts have involved building spatial structure into these dynamics. The addition of a third dynamic variable of empty space to the replicator equation causes its two-phenotype population to have lower density when Defectors predominate, which then favours the production of Cooperators in randomly forming discrete interaction groups [3,57]. Higher population densities resulting from the benefits of elected cooperation can promote spatial heterogeneity in the distribution of Cooperators when Defectors diffuse slowly [58,59]. All of
these models use a positive Reward. They consequently obtain Cooperator and Defector polymorphisms only with weak altruism in SD, or they depend on positive assortment of strategies to sustain net-cost fitness transferrals between social players in PD [20].

Investigations of social dilemmas stand to benefit from the application of game theory to competitive scenarios with negative $R$ and/or negative $P$.

8. CONCLUSIONS

A net-cost transferral of fitness from strong altruist to beneficiary can be sustained without assortative mixing; the only component of altruism that requires assortative mixing is the indirect fitness gain that comes back to the donor as a result of its costly donation to kin or other altruists. This distinction leads to recognition of a new range of states that we call incidental altruism, in which parasitic exploitation by the beneficiary drives the direct fitness cost to the donor while selection on the donor for assortative mixing drives its indirect fitness gain from reciprocation or kin. Game-theoretic analyses have not previously considered this route to the evolution of cooperative behaviour, which is potentially widespread under density-dependent population regulation and has no lower threshold for assortative mixing. It further suggests the possibility that empirical studies interpreting net-cost transferrals of fitness as evidence of altruism may actually be seeing an interaction driven by parasitism (which we call ‘apparent altruism’, unless accompanied by assortative mixing as ‘incidental altruism’). Disaggregating the drivers for direct and indirect fitness consequences requires knowledge of the absolute payoff for the interactions amongst the donor’s beneficiaries. Only if this is non-negative is the net-cost transferral of fitness driven by altruism from the donor.

We thank Miguel Gonzalez Canudas, Simon Powers, Paul Ryan, and two anonymous reviewers for insightful comments.
REFERENCES


Klausen, M., Aaes-Jorgensen, A., Molin, S. & Tolker-Nielsen, T. 2003 Involvement of bacterial migration in the development of complex multicellular structures in 


Press.


Van Dyken, J. D. 2010 The components of kin competition. _Evolution_ **64**, 2840-2854.


Table 1. Matrix of interaction payoffs for the actor in two-strategy games, assuming $R > P$.

<table>
<thead>
<tr>
<th>actor \ recipient</th>
<th>Cooperator</th>
<th>Defector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperator</td>
<td>Reward, $R$</td>
<td>Sucker, $S$</td>
</tr>
<tr>
<td>Defector</td>
<td>Temptation, $T$</td>
<td>Penalty, $P$</td>
</tr>
</tbody>
</table>
Table 2. Specifications for the seven social dilemmas and one HG with net-cost transferrals of fitness from C Cooperators to D Defectors (the +/- interactions T > 0 > S, as stippled in figure 2).

<table>
<thead>
<tr>
<th>Specification*</th>
<th>Game</th>
<th>Outcome</th>
<th>Players</th>
<th>Phenotypes</th>
<th>Populations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Strong altruist to selfish beneficiary (R &gt; P ≥ 0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. T &gt; R &gt; P ≥ 0 &gt; S</td>
<td>PD</td>
<td>D only</td>
<td>figure 2(a)</td>
<td>eqns (4.1), (4.2)</td>
<td>-</td>
</tr>
<tr>
<td>2. R ≥ T &gt; 0, P ≥ 0 &gt; S</td>
<td>SH</td>
<td>C or D</td>
<td>figure 2(a)</td>
<td>eqns (4.1), (4.2)</td>
<td>-</td>
</tr>
<tr>
<td>(b) Apparent or incidental altruist victim to parasitic manipulation (R ≥ 0 &gt; P)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. T &gt; R ≥ 0 &gt; P ≥ S</td>
<td>PD</td>
<td>D only</td>
<td>figure 2(b)</td>
<td>eqns (4.1), (4.2)</td>
<td>-</td>
</tr>
<tr>
<td>4. T &gt; R ≥ 0 &gt; S &gt; P</td>
<td>SD</td>
<td>C and D</td>
<td>figure 2(b)</td>
<td>eqns (4.1), (4.2)</td>
<td>eqn (5.6)</td>
</tr>
<tr>
<td>5. R ≥ T ≥ 0 &gt; S &gt; P</td>
<td>HG</td>
<td>C only</td>
<td>figure 2(b)</td>
<td>eqns (4.1), (4.2)</td>
<td>eqn (5.6)</td>
</tr>
<tr>
<td>6. R ≥ T ≥ 0 &gt; P ≥ S</td>
<td>SH</td>
<td>C or D</td>
<td>figure 2(b)</td>
<td>eqns (4.1), (4.2)</td>
<td>-</td>
</tr>
<tr>
<td>(c) Apparent or incidental altruist victim to parasitic manipulation (0 &gt; R &gt; P)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. T ≥ 0 &gt; R &gt; P ≥ S</td>
<td>PD</td>
<td>D only</td>
<td>figure 2(c)</td>
<td>eqns (4.1), (4.2)</td>
<td>eqn (5.1)</td>
</tr>
<tr>
<td>8. T ≥ 0 &gt; R &gt; S &gt; P</td>
<td>SD</td>
<td>C and D</td>
<td>figure 2(c)</td>
<td>eqns (4.1), (4.2)</td>
<td>eqn (5.1)</td>
</tr>
</tbody>
</table>

*Conditions are those in figure 1(a) of R > P and R > S for a social dilemma, which additionally requires 2R > T + S.
FIGURE LEGENDS

Figure 1. The four canonical games for two strategies on the continuous Temptation-Sucker plane (in gold and blue respectively), calibrated against constant values $R > P$. (a) Conditions on the table-1 payoff matrix for Prisoner’s Dilemma (PD), Snow-Drift (SD), Harmony Game (HG), and Stag Hunt (SH). (b) Game outcomes for Cooperator and Defector strategies.

Figure 2. Game outcomes as figure 1(b), specifically setting $P = \pm 1/R$. Social dilemmas lie below the angled pink line denoting $2R > S + T$, and $T > R$ or $P > S$, given $R > P$; net-cost fitness transfers from Cooperator to Defector occur in the stippled +/- quartile. (a) SH takes the origin of the T-S plane (because $R > P \geq 0$, here with $P = 1/R$ and $R = 5/3$). (b) HG takes the origin of the T-S plane ($R \geq 0 > P$, $P = -1/R$ and $R = 3/5$). (c) SD takes the origin of the T-S plane ($0 > R > P$, $P = 1/R$ and $R = -3/5$). (d) SD takes the origin of the $\alpha$ plane for competing genotypes G1 and G2 ($k_1 > k_2 > 0$, here with $k_2/k_1 = 3/5$).
(a)

- HG: \( R > P \)
  - \( \leq R \)
  - \( > R \)
- SD: \( R > P \)
  - \( \leq R \)
  - \( > R \)

(b)

- C-only
- C + D
- C or D
- D-only