How Threats Influence the Evolutionary Resolution of Within-Group Conflict

Michael A. Cant1,* and Rufus A. Johnstone2

1. Centre for Ecology and Conservation, University of Exeter, Cornwall Campus, Tremough, Penryn, Cornwall TR10 9EZ, United Kingdom; 2. Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, United Kingdom

Submitted May 25, 2008; Accepted December 16, 2008; Electronically published April 17, 2009

Abstract: Most examples of cooperation in nature share a common feature: individuals can interact to produce a productivity benefit or fitness surplus, but there is conflict over how these gains are shared. A central question is how threats to exercise outside options influence the resolution of conflict within such cooperative associations. Here we show how a simple principle from economic bargaining theory, the outside option principle, can help to solve this problem in biological systems. According to this principle, outside options will affect the resolution of conflict only when the payoff of taking up these options exceeds the payoffs individuals can obtain from bargaining or negotiating within the group; otherwise, threats to exercise outside options are not credible and are therefore irrelevant. We show that previous attempts to incorporate outside options in synthetic models of reproductive conflict fail to distinguish between credible and incredible threats, and then we use the outside option principle to develop credible synthetic models in two contexts: reproductive skew and biparental care. A striking prediction of our analysis is that outside options are least relevant to the resolution of conflict in cooperative groups of kin and are most relevant in transient associations or interactions among nonrelatives. Our analysis shows a way to link the resolution of within-group conflict to the environmental setting in which it occurs, and it illuminates the role of threats in the evolution of social behavior.

Keywords: reproductive skew, cooperation, biparental care, mutualism.

Introduction

In many situations in nature, two or more individuals can cooperate to produce mutual fitness benefits, but there exists conflict over how the benefits of cooperation are shared. Cooperatively breeding species, for example, can together produce more young than can solitary breeders, but within groups, individuals compete to monopolize reproduction (Keller and Reeve 1994; Clutton-Brock 1998; Johnstone 2000; Reeve and Keller 2001; Magrath et al. 2004; Cant 2006). In biparental care systems and interspecific mutualisms, individuals can gain from coordinated action, but each would prefer to reap the maximum return benefit from the minimum investment in the interaction (Trivers 1971; Houston and Davies 1985; Bull and Rice 1991; McNamara et al. 1999; West et al. 2002; Sachs et al. 2004; Foster and Wenseleers 2006; Johnstone and Hinde 2006). The way in which such conflicts are resolved is expected to influence individual behavior, group-level characteristics, and the ease with which cooperative associations form. Consequently, understanding the principles of conflict resolution is a major focus of research on social evolution (Alexander 1987; Noé and Hammerstein 1994, 1995; Godfray 1995; Reeve 2000; Frank 2003; Cant 2006; Ratnieks et al. 2006).

An outstanding question is how the availability of outside options, such as leaving to seek partners elsewhere or evicting others from the group, affects the stability of cooperative groups and the evenness of profit sharing. To address this question requires an approach that links within-group conflict to the ecological and social environment in which it occurs. This link is usually missing, however, from traditional “forced play” models of cooperation in which the probability of repeated encounter is assigned exogenously to the model and individuals cannot choose to exit an interaction in response to defection (Axelrod and Hamilton 1981; Nowak and Sigmund 1992, 1993). By contrast, the link to the setting of an interaction is central to models of reproductive skew, which explore conflict over reproduction in cooperative societies (Vehrencamp 1983; Reeve and Ratnieks 1993; Johnstone 2000; Reeve and Shen 2006), and to the theory of biological markets (Noé et al. 1991; Noé and Hammerstein 1994; Bshary and Noé 2003). Reproductive skew models, in particular, have sought to clarify the role of outside options in conflict resolution, and so are a useful starting point in the search for general principles that can be applied to other forms of cooperation such as biparental care and mutualism.

Two main types of skew model have been proposed to
explain variation in the distribution of reproduction within groups. Transactional models (Vehrencamp 1983; Reeve 1991; Johnstone and Cant 1999; Reeve 2000; Buston et al. 2007) assume two-player cooperative associations in which one player (the allocator; Buston et al. 2007) can choose how productivity is shared but the other player can threaten to break up the association (by leaving or evicting its partner) unless offered a large enough share. Transactional models, therefore, assume that the level of skew in a stable group is determined by the inclusive fitness value, to the nonallocating player, of exercising this threat. By contrast, compromise models ignore the possibility of outside options and focus instead on how conflict is resolved when both parties exert partial, costly control over the outcome, for example, via a tug-of-war over reproductive shares (Reeve et al. 1998) or a competition to produce the most young (Cant 1998; Cant and Johnstone 1999). Thus, we have a battleground of conflict (Godfray 1991), the boundaries of which are defined by the payoffs to the two players of pursuing alternative, noncooperative strategies, and we also have some candidate models to explore how conflict within this battleground is resolved (Cant 2006). However, does the location of the boundaries influence the resolution of conflict within them? While there have been several attempts to synthesize skew models in order to answer this question, there is still little agreement as to which approach is appropriate (Johnstone 2000; Reeve and Shen 2006; Nonacs 2007).

A possible solution to this problem can be gleaned from work in economic bargaining theory, which is concerned with problems that are structurally very similar to ours (Nash 1950; Osborne and Rubinstein 1990; Muthoo 1999). In economics, a bargaining situation is one in which two or more players have a common interest in cooperation, but they disagree over how the profits of cooperation should be distributed. An obvious example is trade, although formally similar problems arise in a range of other contexts, from salary negotiations and litigation to military disputes between nation states (Fearon 1995; Powell 1996, 2002; Reiter 2003). There is a large literature on bargaining, which we do not attempt to review here; readers seeking an accessible introduction are directed to Muthoo (1999, 2000). Rather, we borrow a graphical model from this literature and use it to show how a simple principle of bargaining theory, the outside option principle (Binmore 1985; Sutton 1986; Binmore et al. 1989), can help to solve the problem of how to synthesize models of reproductive skew and illuminate the role of threats in the resolution of evolutionary conflict in other contexts such as biparental care and mutualism.

A Graphical Model of Conflict Resolution

Bargaining situations can be represented in a convenient graphical form (Nash 1950, 1953; Osborne and Rubinstein 1990; Powell 2002). This graphical model is a useful tool to help understand and visualize how social conflicts are resolved. In figure 1a, we plot the potential fitness out-

Figure 1: Graphical model of conflict resolution. a, Simple bargaining situation in which neither player has an outside option and negotiation is cost free. The shaded area is the set of possible outcomes from the negotiation. Resolution at point R represents a case where player 2 has greater bargaining power than player 1; resolution at point S represents the case where the reverse is true. b, Example with outside options. Players 1 and 2 have outside options of values $u_1$ and $v_2$, respectively, represented by a single point X. In addition, in this case we assume that the negotiation process is costly, reducing the total fitness that can be realized through cooperation by a factor $c$ ($0 < c < 1$). Where player 2 has greater bargaining power than player 1, the resolution at R results in a higher payoff for both players than their outside option, so in this case neither player’s threat to exercise their outside options is credible. Where player 1 has greater bargaining power, resolution at point S results in a payoff $v'$ for player 2 that is less than their outside option payoff $v_2$, so player 2’s threat to exercise their outside option is credible. In this case, player 1’s best option is to allow a shift along the line of efficient outcomes to point $S'$, at which point player 2’s threat is no longer credible.
comes of an interaction between two individuals—player 1 and player 2—who can cooperate in some way to produce a fitness surplus. We will assume for simplicity’s sake that the two players are unrelated, and we plot the direct fitness payoffs of the two players along orthogonal axes denoted \( u \) and \( v \), respectively. The set of possible direct fitness outcomes from the interaction is shown as the shaded region in figure 1a. Note that the model can be applied equally well to interactions among individuals of symmetrical or asymmetrical relatedness: in this case, we would plot inclusive fitness, rather than direct fitness, along the orthogonal axes, and the shaded area would represent the paired inclusive fitness outcomes of an interaction. The graphical approach might also be extended to three players by adding a third axis, the \( z \)-axis; the set of possible fitness outcomes would then be represented as a solid in three dimensions.

If the players choose not to cooperate, or if they cannot agree on coordinated action, we assume that they both receive a fitness payoff of 0. We assume in all figures that the maximum total fitness (or surplus) achievable by interaction is equal to a constant \( k \), so that the boundary of the set fitness outcomes is linear. The boundary defines “efficient” outcomes, that is, outcomes at which a player can increase its own payoff only at the expense of the other player. We define a conflict as resolved when neither player can gain from further attempts to influence the outcome; that is, the outcome is efficient and in equilibrium.

The expected efficient resolution within the set of possible fitness outcomes depends on the rules governing the bargaining or negotiation process and the relative attributes of the players. Let us assume initially that the negotiation involves no personal or group costs. If player 1 can make offers of a division of the surplus that player 2 can only accept or reject, then the problem takes the form of an “ultimatum game” (Guth et al. 1982). Accordingly, we expect the resolution at \( A \); that is, we expect player 1 to claim all of the available fitness (i.e., \( k \) units) and leave player 2 with a share that is infinitesimally greater than 0 (or with the smallest allocatable unit of fitness). Likewise, if player 2 makes offers that player 1 can only accept or reject, we expect a resolution at \( B \). The two resolution points at \( R \) and \( S \) correspond to cases where both players can exert partial influence over the outcome, but the players differ in some way in their “bargaining power,” that is, in their effectiveness at negotiating toward their favored resolution. Resolution \( R \) (at which the two players receive payoffs \( u_1 \) and \( v_1 \), respectively) represents the case when player 2 exerts greater influence over the outcome of bargaining than player 1; \( S \) represents the resolution when the opposite is true. In biological systems, differences in bargaining power will arise when players differ in the fitness costs they experience during negotiation (e.g., when this involves physical risk or display costs) or in their sensitivities to delay (e.g., where there are time constraints to realize the gains of cooperation). Bargaining power is therefore a relative measure of the fitness costs paid to achieve a given share of the productivity surplus of the association. Where interactants are related, the two players’ inclusive fitness costs of bargaining will be correlated, and increasing relatedness is expected to reduce the impact of underlying asymmetries in bargaining power. Where relatedness is asymmetrical (e.g., in social Hymenoptera), the less related party will possess greater effective bargaining power (other things being equal) because it is less sensitive to the costs it inflicts on the other player.

The Outside Option Principle

Our aim is to determine how the presence of outside options influences the resolution of conflict. Figure 1b suggests a simple solution to this problem. In this figure, we again consider two cases that depend on whether player 1 or player 2 is the stronger bargainer. Additionally, in this figure we assume that negotiation is costly in that some of the realizable fitness surplus is used up in the negotiating process; specifically, during negotiation, maximum total fitness is reduced by some constant factor \( c \); in reality, \( c \) may itself vary with the characteristics of the bargainers, such as relative bargaining power or relatedness. The outside options available to the players are represented by a single point \( X \), which gives the fitness payoffs of the two players if either one chooses to break up the association.

When player 2 is stronger than player 1, the process of negotiation leads to an efficient resolution at point \( R \). At this resolution point, both players receive a greater payoff than they would get if either was to exercise the outside option. Thus, in this case, a threat by either party to exert the outside option is not credible in the sense that, having reached this point in the game, it would not pay for them to do so (Nash 1953; McNamara and Houston 2002). Consequently, the presence of the outside option makes no difference to the resolution of the conflict.

In the case where player 1 is stronger than player 2, negotiation leads to an efficient resolution at point \( S \). However, player 2’s payoff at point \( S \) is lower than its outside option payoff, so player 2’s threat to exert its outside option is credible. Deterring player 2 from exercising this outside option is very much in player 1’s interest (because \( u^* > u_2 \)), so player 1 will gain from permitting a shift in the division of the surplus along the line of efficiency to point \( S^* \), which yields a payoff for player 2 that is just equal to its outside option. The outside option in this case clearly does affect the outcome of the negotiation process.

In summary, the graphical model suggests that a focal individual’s outside option will influence the resolution of
within-group conflict, but only if this option yields a greater payoff to the focal individual than the payoff it expects to receive through negotiation (Nash 1953; Binmore 1985; Binmore et al. 1989). In these circumstances, the focal individual can use a credible threat to exercise its outside option to obtain a more favorable resolution, one that is sufficient to render the threat incredible (this conclusion was also reached independently by Buston and Zink [2009]). In economics, this simple result has been called the outside option principle (Sutton 1986; Muthoo 1999, 2000). In pairwise interactions, only one player’s threat to exercise an outside option can be credible at a time. In larger groups, however, two or more players may simultaneously exercise credible threats to extract a greater share of resources from the others (Raiffa 2005).

Previous Attempts at Synthesis

How does this analysis compare with previous attempts to determine the role of outside options in the resolution of evolutionary conflict? There have been two such attempts to produce a synthetic model of within-group conflict, each in the context of reproductive skew.

Johnstone (2000) first attempted to incorporate outside options and incomplete control (or bargaining) in a single framework. He considered conflict over reproduction between a dominant individual and a subordinate individual in a cooperative association who were symmetrically related by coefficient $r$ (the same analysis can be applied to cases of asymmetric relatedness, although this was not explored in Johnstone’s [2000] article). In this model, the two individuals can together generate a maximum productivity $k$ relative to a lone breeder (for concreteness, one can think of $k$ as the number of surviving offspring produced per time unit). The dominant and subordinate both possess outside options such that they require fractions $Q$ and $P$, respectively, of this total productivity to make staying in the group worthwhile. In the first step of the game, the two players concede these fractions to each other so that the association is stable. They then compete for the remaining contested productivity via a costly tug-of-war in which both players invest selfish effort to obtain a greater share of it at a cost to the total productivity to be shared. A graphical representation of the model is shown in figure 2a.

A problem with this approach is that the minimum fractions $P$ and $Q$ that are set aside in the first step in order to guarantee group stability are calculated on the basis of total productivity $k$, overlooking the fact that some of this productivity will be used up in the costly tug-of-war in the second step. After a tug-of-war, these shares are no longer sufficient to ensure group stability. The model of Reeve and Shen (2006) attempted to address this problem using an alternative approach. Again, they assumed that two players, a dominant and a subordinate, engage in a game in which each can choose two quantities: the fraction they each set aside as a concession to the other (denoted $P$ and $Q$, respectively) and the effort they expend in a tug-of-war over the remainder (denoted $x$ and $y$, respectively). However, in this case, the authors assume that an evolutionarily stable outcome is reached when two simultaneous conditions hold: (1) each player receives an inclusive fitness payoff equal to the player’s outside option and (2) each player’s selfish effort is a best response to the other’s (i.e., selfish efforts are at a Nash equilibrium). These two conditions can hold only at a single point in the set of possible outcomes represented...
in figure 2b; the point represented by the outside option point X (if the players are related, the point X represents the single point in inclusive fitness space of the two players). Consequently, at the solution obtained by Reeve and Shen (2006), the players use up all of the potential benefits of group membership in the process of finding a resolution (Nonacs 2007). The two solution conditions of Reeve and Shen’s (2006) model force a resolution at which each player gives away a free share of reproduction to their opponent but at the same time fights hard for the remainder. Intuitively, it would make more sense for the players to reduce the size of the free share offered to their partner and invest less effort in wasteful competition. This intuition is borne out; in the appendix, we show that no stable solution is possible when a player offers a concession while simultaneously fighting over the contested share of reproduction.

A similar conclusion has been reached by authors modeling within-group conflict in other contexts. For example, Gardner et al. (2007) developed a model of sex ratio conflict in parasitoid polyembryonic wasps in which (haploid) males prefer a less female-biased sex ratio than their diploid sisters. In these species, a fraction of larvae develop precociously as soldiers, giving up their own future reproduction in order to kill competitors (including opposite sex siblings) developing in the same host (Gardner et al. 2007). In this case (and in the begging models of Godfray 1991, 1995), the battleground of conflict is defined by the males’ and females’ different optima (i.e., by optimization constraints, as opposed to by the group stability constraints that define the battleground in transactional skew models; Cant 2006). Male offspring can allocate a free handout to their sisters (by limiting their clonal proliferation in a host) or engage in costly competition akin to the tug-of-war (by increasing the proportion of clones that develop as soldiers). Gardner et al. (2007) show that male offspring are never favored to offer a free handout to females while at the same time engaging in costly conflict, as is expected from the proof we present in the appendix.

Both previous attempts to produce a synthetic model of reproductive skew, therefore, face strong theoretical objections. These models have also been forcefully critiqued by Nonacs (2007), who highlights some of the same problems pointed out here (however, the simulation model offered by Nonacs [2007] does not include outside options and so does not solve the problem of how to synthesize skew models). In the next section, we show how the outside option principle helps to resolve the problems faced by previous models and illustrate the wider applicability of this principle by outlining credible synthetic models of within-group conflict in two contexts: conflict over reproduction in cooperative breeders and conflict over investment in young in biparental care systems.

### Credible Synthetic Models

#### Reproductive Skew

How can we construct a synthetic model of reproductive skew that corresponds to the graphical model shown in figure 1b and consequently avoids the problems encountered by previous analyses? Like previous analyses, we will consider groups composed of a dominant individual and a subordinate individual who are symmetrically related by r. For simplicity’s sake, we assume that the dominant’s only outside option is to evict the subordinate (at no cost); in this case, the dominant’s direct fitness payoff as a sole breeder on the patch is standardized at 1 (it would be simple to incorporate a cost of eviction or to assume that dominants could also choose to leave). The subordinate’s outside option is to disperse to breed independently; from this, it can expect direct fitness payoff X (<1). The two players engage in a tug-of-war over reproduction in the group, but they adjust their efforts if outside options become relevant during the bargaining process.

In the tug-of-war game, the dominant can invest any fraction x and the subordinate can invest any fraction y of the remaining group productivity in an effort to maximize their inclusive fitness. The dominant thereby obtains a share x/(x + y) and the subordinate obtains a share y/(x + y) of total productivity k(1 − x − y). The inclusive fitness payoffs of the dominant and subordinate from a tug-of-war in the absence of outside options are, respectively:

\[
I_d = k(1 - x - y) \left( \frac{x}{x + y} + \frac{r \cdot y}{x + y} \right) \tag{1}
\]

\[
I_s = k(1 - x - y) \left( \frac{y}{x + y} + \frac{r \cdot x}{x + y} \right) \tag{2}
\]

where the parameter b measures the relative strength or efficiency of the subordinate.

For any given effort x by the dominant, there is a best response function for the subordinate, y*(x), which is found by solving ∂I_d/∂y = 0. Similarly, for any effort y by the subordinate, there is a best response function for the dominant, x*(y), which is found by solving ∂I_s/∂x = 0. The stable Nash equilibrium solution (which we will denote (x*, y*)) is where the players’ effort levels are best responses to each other; that is, x*(y) = y*(x). At this equilibrium, the dominant and subordinate receive payoffs I_d* = I_d(x*, y*) and I_s* = I_s(x*, y*), respectively. Whether the presence of outside options will affect the outcome of bargaining depends on the magnitude of I_d*, I_s* and the value of the subordinate’s outside option X. There are four possibilities, which we consider in turn.
No Concessions. If both players obtain greater payoffs from the tug-of-war than they can expect to get from pursuing their outside options, the outside options are irrelevant. That is, where

\[ I_d(x^*, y^*) > 1 + rX, \]  
\[ I_s(x^*, y^*) > X + r, \]  
the outcome is that of the simple tug-of-war. We label this zone of parameter space “no concessions” because neither player needs to take into account the outside options of the other when deciding on its own optimal effort in the tug-of-war.

Dominant Yields a Concession. Where

\[ I_d(x^*, y^*) > 1 + rX, \]  
\[ I_s(x^*, y^*) < X + r, \]  
the subordinate will favor dispersal over staying, whereas the dominant will favor maintaining the association. In this region, the dominant must ease off in the tug-of-war (what we will term “granting a concession”) if it is to retain the subordinate in the group. Specifically, the dominant should invest at a level \( x' \) such that

\[ I_s(x', \hat{y}(x')) = X + r. \]

To find the stable best efforts in this region, we set expression (2) equal to \( X + r \), solve for \( x \) to yield a function \( x' \), and substitute this in place of \( x \) into our expression for \( \hat{y}(x) \). After rearranging, we can solve for \( y \) to find the best response effort of the subordinate \( \hat{y}(x') \), given that the dominant invests effort \( x' \). The dominant will favor investing at the lower level \( x' \) so long as

\[ I_d(x', \hat{y}(x')) > 1 + rX. \]  

If this inequality is not satisfied, the group will break up because the dominant will not be selected to reduce its competitive effort in the tug-of-war and the subordinate will disperse. Thus, the “dominant yields a concession” zone is defined as that for which conditions (5), (6), and (7) hold.

Subordinate Yields a Concession. Where

\[ I_d(x^*, y^*) < 1 + rX, \]  
\[ I_s(x^*, y^*) > X + r, \]  
the dominant will favor evicting the subordinate and the subordinate will favor remaining in the group. To avoid eviction, the subordinate must therefore ease off in the tug-of-war. Specifically, the subordinate must invest at level \( y' \), which satisfies

\[ I_s(\hat{x}(y'), y') = X + r, \]  

where \( \hat{x}(y') \) is the best response of the dominant when the subordinate invests at level \( y' \). The subordinate will favor investing at this lower level \( y' \) so long as

\[ I_d(\hat{x}(y'), y') > X + r. \]  

If condition (10) does not hold, the subordinate will not be favored to ease off and so will be evicted (because we assume no cost of eviction, this would be observationally indistinguishable from voluntary dispersal by the subordinate). Together, conditions (8), (9), and (10) define the “subordinate yields a concession” (or subordinate restraint) zone.

Group Breaks Up. Finally, as described, where both players’ outside options exceed their payoff from the tug-of-war (i.e., neither condition [3] nor condition [4] holds) or where one player requires a concession to favor keeping the group intact but it does not pay the other individual to grant this concession (i.e., conditions [5] and [6] hold but not condition [7], or conditions [8] and [9] hold but not condition [10]), then the subordinate will leave or be thrown out and the group will break up.

We have derived analytical solutions for the boundaries of the regions of parameter space that correspond to the four outcomes described above. Because these expressions are in some cases complex, we will not give them here; instead, figure 3 graphically shows these regions of parameter space for the cases of unrelated \( (r = 0) \) and related \( (r = 0.5) \) competitors. The general qualitative pattern is simple: groups are more likely to break up when productivity \( (k) \) is low, the subordinate’s opportunity for independent breeding \( (X) \) is great, and competitors are less closely related. There is a substantial region in which both the threat of departure and the threat of eviction prove to be incredible, so that the solution of the basic tug-of-war game is stable and unaffected by outside options. This outcome is most likely when group productivity is high, there is little opportunity for independent breeding, and competitors are more closely related. When the subordinate is much weaker than the dominant, and particularly when there are substantial opportunities for independent breeding and the competitors are unrelated, the threat of subordinate departure becomes credible and the dominant must ease off in competition to retain the subordinate in the group. Conversely, when the subordinate is not too
much weaker than the dominant, when there are substantial opportunities for independent breeding, and when the competitors are unrelated, the subordinate may have to ease off in competition for its presence to be tolerated.

**Biparental Care**

The same method can be used to incorporate outside options into models of evolutionary conflict in other contexts, for example, biparental care systems of the kind modeled by Houston and Davies (1985) and McNamara et al. (1999, 2003).

Consider two parents, a male and a female, who can choose to invest in their brood at levels $m$ and $f$, at a cost to their future fitness of $C_m(m)$ and $C_f(f)$, respectively (where $\partial C_m/\partial m, \partial C_f/\partial f > 0$; $\partial^2 C_m/\partial m^2, \partial^2 C_f/\partial f^2 \geq 0$). We assume for simplicity’s sake that there is no extra-pair paternity. The total number of surviving young is an increasing, decelerating function of total investment $t$ ($t = m + f; \partial B/\partial t > 0, \partial^2 B/\partial t^2 < 0$). Thus, male and female fitness is given by

$$W_m(m, f) = \frac{B(m + f)}{2} - C_m(m), \quad (11)$$
Figure 4: Biparental care when males may exert a credible threat to desert. a, Zones of stability as a function of the relative cost of male care and the value of the male’s outside option. Results are shown for a model in which male and female fitness functions take a simple quadratic form (specifically, $B = (m + f) - (m + f)^2$; $C_m = c m^2$; $C_f = f^2$). b, Evolutionarily stable levels of parental effort for males and females for a single slice through this parameter space (shown as a short vertical line in a). In this example, we assume that investment is costlier for females, so females initially invest at lower levels than males ($f < m$). However, where the male’s threat to depart is credible, the female may be selected to work harder (at rate $f'$) so that the male’s optimum effort is lowered (to $m'$) and he continues to favor staying in the partnership. However, where the male’s outside option is more profitable, the extra investment required to keep the male in the partnership is not worth it for the female, so she favors allowing the male to leave and switching to the optimum work rate as the sole carer ($f_0'$).

$$W(m, f) = \frac{B(m + f)}{2} - C_f(f).$$

For any given female effort $f$, there will be a best effort for the male to invest $\hat{m}(f)$, which is found by solving $\partial W_m / \partial m = 0$. Similarly, for any given male effort $m$, there will be a best effort for the female, which is found by solving $\partial W_f / \partial f = 0$. If the fitness functions (11) and (12) are such that both $\partial \hat{m}(f) / \partial f$ and $\partial f(m) / \partial m$ are negative and of magnitude $< 1$, there may exist a stable Nash equilibrium at which both parents invest nonzero effort ($m^*, f^*$).

To introduce outside options in a simple way, we will assume that only males have the option to desert to look for breeding opportunities elsewhere (although the analysis could easily be extended to allow both parties the opportunity to desert). If the male deserts, he can gain the expected payoff $X_m$ elsewhere (reflecting mating opportunities outside the partnership), while the female gains the expected payoff

$$X_f = \frac{B(f'\star)}{2} - C_f(f_\star\star),$$

where $f_\star\star$ is the optimum investment of the female given $m = 0$ (i.e., the male deserts). In the region where $W_m(m\star, f\star) > X_m$, the presence of the desertion option does not influence the best efforts of the male and female. However, if $W_m(m\star, f\star) < X_m$, the female needs to work at a higher rate $f'$ to retain the male, such that $W_m(m(f'), f') = X_m$.

The female will favor working at this higher rate so long as $W_m(m(f'), f') > X_f$.

If this inequality is not satisfied, the female will favor switching to working at the lone breeder optimum rate, $f_0\star$, forcing the male to disperse.

We have obtained analytical results for this model assuming that the fitness functions take a simple quadratic form. The results are presented graphically in figure 4a, which shows the three zones where the male’s threat to leave is incredible, where it is credible, and where the partnership breaks up. Figure 4b shows the levels of investment for a particular value of the males outside option $X_m$. Where the outside option of the male is of relatively low value, the male’s threat to leave is incredible and so it exerts no influence on the evolutionarily stable levels of investment for the male and female. When the outside option of the male is of high value, the female does her best to allow the male to depart and to raise the brood on her own rather than working harder to induce the male...
to stay. In between, there is a region for which the male’s threat to leave is credible and the female does her best to work harder to keep the male in the group. Interestingly, in some cases, the female ends up working harder than she would if she were raising the brood on her own. This result arises because in a partnership, the female can share the costs of total investment with the male even if she ends up paying a disproportionate share of these costs.

Discussion

Outside Options and Conflict Resolution

The question we have considered is whether the presence of outside options should influence the resolution of conflict in cooperative associations. The answer according to our analysis is yes, but only under certain circumstances. Outside options are irrelevant when negotiation yields a higher payoff to both players than they could expect from pursuing their outside option. The role of outside options, therefore, depends on the balance between the payoffs obtained by negotiation and those obtained by carrying out a threat.

A striking prediction of our analysis of reproductive conflict is that outside options will be least relevant where group members are close relatives and the productivity benefits of association are high: precisely, those situations where selection is most likely to favor cooperation in the first place (e.g., in cooperatively breeding insects and vertebrates) and to which reproductive skew models are most often applied. Threats to exercise outside options are less likely to be credible in these circumstances because alternatives outside the group are less profitable. This may explain why the two studies of cooperative breeders that manipulated ecological constraints experimentally (Langer et al. 2004; Heg et al. 2006) found no effect on reproductive sharing. Outside options should play a more important role where individuals are not tightly bound by kinship, where the productivity benefits of association are low, and where it is relatively easy to recruit alternative partners. These conditions may apply best to social groups in which group members contribute little to productivity (e.g., coral-dwelling fish; Buston 2003; Wong et al. 2007), to facultative or short-term cooperative associations (e.g., coalitions or alliances: Russell 1983; Zabel et al. 1992; Connor et al. 1999; Silk et al. 2004; grooming relationships: Barrett et al. 1999; Schino 2007), and to some interspecific mutualisms (e.g., Bshary and Grutter 2005). Consequently, we expect manipulation of outside options to have a greater effect on the partitioning of productivity (and levels of cooperation) in these contexts.

What decides the payoffs of bargaining versus exercising an outside option? In a biological context, bargaining or negotiation can be thought of as any interaction or behavioral process by which animals with conflicting interests can reach a compromise settlement (or stable behavioral equilibrium) without breaking up the association or resorting to an all-out fight. This may involve the exchange of signals or displays, physical aggression, or other attempts to control resources. In many cases, the details of the negotiation process will be unclear. Nevertheless, we can expect an individual’s bargaining power to depend (as it does in economics) on its ability to sustain costs or inflict them on its partner during the negotiation process and on its sensitivity to delay in reaching a settlement. Bargaining power in nature, therefore, will typically correlate with relative quality, fighting ability or strength, or access to group resources. Bargaining power is not expected to depend on the availability of options outside the group.

By contrast, the payoffs associated with outside options will depend primarily on what can be termed “market forces”—the quality and availability of alternative partners or vacancies in the wider population—rather than on individual attributes of the bargainers (although these attributes may also affect the value of outside options; this is especially likely in the special case where the outside option is to enter into an escalated fight). Where individuals can readily terminate an interaction and recruit a new partner, threats of departure and eviction will often be credible and so may influence the outcome of negotiation in a way that overrides differences in relative bargaining power. The power of such sanctions has been the focus of much discussion in the literature on mutualisms (Herre et al. 1999; Denison 2000; Ferriere et al. 2002; West et al. 2002; Kiers et al. 2003; Sachs et al. 2004; Foster and Wenseleers 2006) and biological markets (Noé and Hammerstein 1994, 1995; Noé 2001; Johnstone and Bshary 2002; Bshary and Noé 2003). These models suggest that the level of cooperation will depend on the quality and availability of partners (partner choice) and the ability to coerce or compete within partnerships (partner control; Noé and Hammerstein 1994), which is equivalent to the distinction in our analysis between resolution according to outside options versus bargaining power. The outside option principle thus suggests a simple rule to integrate these two types of model: partner choice will affect partner control if and only if the threat to swap partners is credible.

Signaling of Threats

The outside option principle is useful in clarifying the way threats can be used to influence social behavior. In our analysis, we have focused on the threat of departure or eviction from the group, but the same principle holds for other types of threat such as that of escalated fighting.
(another type of outside option). In a cooperative association, each player has a strong incentive to avoid triggering a threat because doing so entails a sudden drop in their own fitness. A threat threshold represents a fitness “cliff edge” for the recipient of the threat (Kokko 2003), so we should expect strong selection on organisms to exercise restraint to avoid triggering threats. In our models, this presents no difficulty because both parties know where threat thresholds lie and whether threats are credible. In reality, however, there will almost always be some uncertainty about the availability of outside options and the costs and benefits of triggering a threat. Threats are credible only if they are perceived as such by the other player, so information is a crucial determinant of whether threats will influence social behavior. How might information about threats be acquired or conveyed?

In repeated interactions (with the same or multiple partners), individuals might gain information about the location of threat thresholds and the credibility of threats through trial-and-error learning (Clutton-Brock and Parker 1995). Alternatively, group members could communicate so as to avoid triggering a threat. However, here there is considerable scope for conflict and deception because an exploiting individual will benefit from knowing exactly where their partner’s threat threshold lies, while the partner will benefit from misrepresenting or concealing the location of their threshold. For example, it may pay subordinate individuals in a dominance hierarchy to conceal their ability or motivation to challenge those of higher rank. In the case of dominance interactions of Polistes wasps, subordinates freeze their position and lower their head and antennae as dominants antennate and mount them (Gamboa et al. 1990). Such submissive behavior probably yields little information about the subordinate’s fighting ability or aggressive motivation. By contrast, in the cooperatively breeding cichlid Neolamprologus pulcher, subordinate individuals engage in energetically costly submissive behavior toward dominants that might yield reliable information about their relative fighting ability or quality (Grantner and Taborsky 1998; see also Kutsukake and Clutton-Brock 2006 for an example of active submission in meerkats). A better understanding of threat signaling might therefore cast light on the variety of dominance-submission behaviors that are a conspicuous feature of animal dominance hierarchies.

Where it pays a threatening individual to signal the location of its threat threshold, there is a clear incentive for deception because signalers gain from exaggerating their willingness to exercise a threat. Signals aimed at conveying an imminent threat will therefore lack credibility unless they are costly in some way (see Fearon 1995). A willingness to enter into an escalated fight might be signaled via provocative or costly acts of aggression. A willingness to permanently evict an individual might be signaled by attacking or chasing the individual or excluding it temporarily from the group (e.g., Mulder and Langmore 1993; Clutton-Brock et al. 1998). It is possible, therefore, that much of the low-level aggression observed in animal societies functions to support the credibility of threats of more extreme action (such as departure, eviction, or escalated fighting). This hypothesis predicts that the level of aggression will vary with the value of outside options to the two players. By contrast, the hypothesis that aggression is the means by which individuals negotiate or bargain predicts that changing the value of outside options will have no influence on aggression levels within the group. Studies to add or remove outside options experimentally (such as those of Langer et al. [2004] and Heg et al. [2006], neither of which focuses on aggression) or to change the expected payoff of all-out conflict would help to distinguish these hypotheses and so elucidate the function of social aggression.

Studying threats empirically is challenging because the most effective threats are those that rarely need to be exercised (e.g., consider the cost-effectiveness of a Mafia protection racket; Gambetta 1994) and because, by definition, the only observed cooperative associations will be those in which group members have avoided triggering group breakup. Revealing latent threats in stable groups may therefore require innovative experimental approaches to disturb the status quo (e.g., Wright and Cuthill 1999; Reeve and Nonacs 1992; Hinde 2006; Wong et al. 2007) or to alter the payoffs of pursuing the outside option (e.g., Koenig 1990; Langer et al. 2004; Cant et al. 2006; Heg et al. 2006). The potential reward is a much-improved understanding of the forces shaping individual behavior in cooperative associations and the conditions for which such associations can form.

Acknowledgments

We thank P. Buston, J. Field, P. Nonacs, K. Reeve, S. Shen, and M. Taborsky for useful discussions and A. Gardner and an anonymous reviewer for comments on the manuscript. M.A.C. was funded by a Royal Society University Research Fellowship.
APPENDIX

Instability of Concessions in the Bordered Tug-of-War

Suppose that the subordinate in a group plays \((Q, y)\). For any level of dominant competitive effort \(x\), one can then solve for the minimum concession sufficient to induce the subordinate to stay, denoted \(\hat{P}(x)\). Following Reeve and Shen (2006), this value satisfies

\[
G(1 - x - y)\left[\hat{P}(x) + (1 - \hat{P}(x) - Q)\frac{by}{x + by}\right] + rG(1 - x - y)\left[Q + (1 - \hat{P}(x) - Q)\frac{x}{x + by}\right] = S + rD, \tag{A1}
\]

where \(G\) is group productivity, \(b\) is the relative strength of the subordinate in a tug-of-war, \(r\) is genetic relatedness, and \(S\) and \(D\) are the direct fitness payoffs of the subordinate and the dominant, respectively, if the subordinate chooses to disperse. The left-hand side of equation (A1) gives the inclusive fitness payoff to the subordinate of remaining in the group; the right-hand side gives the payoff if the subordinate chooses to leave. Equation (A1) yields, after some rearrangement, the following expression for \(\hat{P}(x)\):

\[
\hat{P}(x) = \frac{rD(x + by) + S(x + by) - G(1 - x - y)[b(1 - Q)y + r(x + bQy)]}{G(1 - r)x(1 - x - y)}. \tag{A2}
\]

Note that \(\hat{P}(x) > 0\), implying that the more strongly the dominant competes, the larger the minimum concession it must provide (or, conversely, the less strongly the dominant competes, the smaller the concession it need supply).

In any evolutionarily stable association, the dominant must play a strategy of the form \((\hat{P}(x), x)\), because if \(P < \hat{P}(x)\), it is possible for the dominant to claim a larger share of reproduction without triggering subordinate departure, whereas if \(P > \hat{P}(x)\), the subordinate will leave. Consider, then, the payoff to a dominant that plays \((\hat{P}(x), x)\). The dominant’s inclusive fitness is given by

\[
I = G(1 - x - y)\left[Q + (1 - \hat{P}(x) - Q)\frac{x}{x + by}\right] + rG(1 - x - y)\left[\hat{P}(x) + (1 - \hat{P}(x) - Q)\frac{by}{x + by}\right], \tag{A3}
\]

which, with equation (A2), yields

\[
I = rD + S - G(1 - r)(1 - x - y). \tag{A4}
\]

It follows that

\[
\frac{\partial I}{\partial x} = -G(1 - r) < 0, \tag{A5}
\]

implying that if \(x\) and \(\hat{P}(x)\) are both positive, the dominant would always do better to reduce its competitive effort and offer a correspondingly smaller concession. Intuitively, it makes little sense to offer your competitor a large free concession \(P\) while simultaneously fighting hard over the contested fraction of reproduction. It is more efficient to reduce the size of the concession and compete less strongly; in this way, the competitor’s outside option can be matched (and hence group stability is maintained) at a lower cost to group productivity. An equivalent argument applies to the subordinate (or whichever player is supposed to offer an incentive in the one-sided, bordered-tug-of-war model).
Literature Cited


Associate Editor: Ross H. Crozier
Editor: Michael C. Whitlock