Breaking up fights between others: a model of intervention behaviour

Lee Alan Dugatkin

Department of Biology, University of Louisville, Louisville, KY 40292, USA (laduga01@homer.louisville.edu)

To examine when and why animals break up fights between others in their group, I modelled whether 'winner' and 'loser' effects might be one element driving the evolution of intervention behaviour. I considered one particular type of intervention: when the intervener simply breaks up fights between two others, but does not favour either party in so doing. When victories at time $T + 1$ are more likely given a victory at time T (i.e. winner effects), intervention is often favoured. Intervention is favoured in these circumstances because the intervening party in essence stops others from 'getting on a roll' and climbing up any hierarchy that exists. However, when loser effects alone are at work (defeats at time $T + 1$ are more likely given a defeat at time T), breaking up fights between others is never selected. If both winner and loser effects are operating simultaneously, then the likelihood of intervention behaviour evolving is a function of the relative strength of these two effects. The greater the winner effect relative to the loser effect, the more likely intervention behaviour is to evolve.

Keywords: intervention; aggression; hierarchy; alliances; coalitions

1. INTRODUCTION

A number of factors can determine the outcome of dyadic aggressive interactions between group members. The most obvious of these factors is resource holding power (RHP; Parker 1974; Maynard Smith & Parker 1976; Parker & Rubinstein 1981) of combatants, often measured by some size index. In addition to 'intrinsic' factors such as RHP, 'extrinsic' forces also affect the outcome of aggressive interactions. The three most well-studied extrinsic forces are territory ownership (which typically increases an animal's probability of defeating an opponent; see Pusey & Packer (1995) for a review), winner effects, and loser effects. Winner effects refer to the increased probability of victory at time $T + 1$, as a result of victory at time T . Loser effects refer to the increased probability of being defeated at time $T + 1$, given a prior defeat at time T (see Chase *et al.* (1994) for a review of studies on winner and loser effects, and Landau (1951a,b) and Dugatkin (1997) for theoretical treatments).

In addition to experimentally manipulating extrinsic and intrinsic factors to determine their importance for dyadic aggressive interactions, recent work has shown that much aggression occurs within the context of polyadic interactions (i.e. interactions involving more than two individuals). Examples of coalitions and alliances have been documented in a number of primate species (Harcourt & de Waal 1992; Chapais 1995), but are not restricted to this taxon (Harcourt 1992). For example, coalitions or alliances have been documented in hyenas (*Crocuta crocuta*; Zabel *et al.* 1992), wolves (*Canis lupus*; Fentress *et al.* 1986), lions (*Panthera leo*; Packer & Pusey 1982), cheetahs (*Acinonyx jubatus*; Caro 1994), coatis (*Nasura narica*; Russell 1983) and dolphins (*Tursiops truncatus*; Conner 1992; Connor *et al.* 1992). In most instances, such

polyadic aggressive interactions involve an animal *intervening* in a dyadic interaction between other group members (Harcourt & de Waal 1992). The intervening animal is often dominant in rank to others involved in the interaction, but this need not be the case (see Silk (1992) for a review of rank and intervention). In addition, in primates, interventions most often take the form of the intervener coming to the aid of one of the two other individuals involved in the interaction (Harcourt & de Waal 1992). In principle, however, intervening animals may simply break up an interaction between two others, without siding with either putative combatant. This type of intervention has been found in various primate species (Harcourt & de Waal 1992), and in *Melanochromis auratus*, a cichlid fish (Nelissen 1985).

Various functions have been proposed for intervention behaviour: e.g. supporting kin, reducing 'social disruption', formation of alliances, promotion of group cohesion and 'tension reduction' (see Ehardt & Bernstein (1992) for a review). However, to date it appears that no theoretical work (i.e. mathematical models) has examined the evolution of intervention behaviour. Here, I develop a simple model for the evolution of interventions that serve only to break up fights (the intervener does not side with either party). I hypothesize that intervention is favoured in systems in which winner effects, rather than loser effects, predominate. In addition, I hypothesize that within systems wherein winner effects are important, individuals further up in a hierarchy should be more likely to be involved in breaking up fights between other group members.

2. THE MODEL

For analytical tractability, I will consider the simplest possible group structure that allows for intervention

behaviour: a group of three individuals. Model I will examine winner effects alone, model II will look at loser effects alone, and model III will examine intervention when winner and loser effects operate simultaneously.

I will make the following four simplifying assumptions: (i) dyadic interactions are formed at random within groups; (ii) interventions take place after an initial round of fights that determine the rank of the three group members; (iii) interventions are always successful in the sense that they stop others from undertaking aggressive interaction; and (iv) winner and loser effects have a memory window of one move. For example, a victory during aggressive encounter T would make winning more likely in the next encounter a player had (encounter $T + 1$), but not in encounters $T + 2$, etc.

The following variables are defined:

p_{ij} is the probability that i defeats j , given that no winner or loser effects are in operation, where i and j can assume any of the three possible positions in a hierarchy: α , β or γ .

B_i are the benefits accrued to individual i for its position in the hierarchy (see Pusey & Packer (1995) for a review of such benefits).

C_i are the costs accrued to individual i for intervening in a fight between others. These costs might include risk of injury, energy expenditure, etc. (Silk 1992).

W is the 'winner effect'. This measures an individual's increased probability of victory at time $T + 1$, given prior victory at time T . For analytical reasons, W is represented as the reduced probability of defeating an individual with prior winning experience. $0 \leq W \leq 1$.

L is the 'loser effect'. This measures an individual's increased probability of losing at time $T + 1$, given prior defeats at time T . For analytical reasons, L is represented as the increased probability of defeating an individual with prior losing experience. $0 \leq L \leq (1 - p_{ij})/p_{ij}$. This upper boundary assures that the greatest chance of defeating an opponent with prior defeats is 1.

(a) Model I: winner effects only

(i) Case I: when should α individuals intervene?

Should an α individual opt to intervene in a fight between β and γ , its incremental change in fitness after such an act would be:

$$\frac{p_{\alpha\beta}B_\alpha}{2} + \frac{p_{\alpha\gamma}B_\alpha}{2} - C_\alpha, \quad (1)$$

where the first term is the probability of α defeating β , multiplied by the pay-off for such a victory, weighted by the probability of such an interaction. The second term is the probability of α defeating γ , multiplied by the pay-off for such a victory, weighted by the probability of such an interaction, and the last term is the cost of intervening. Notice that when intervention occurs there are no winner effects in expression (1), as intervention by α does not allow either β or γ to emerge as a victor in their dyadic interaction.

Expression (1) can be rewritten as:

$$\frac{B_\alpha(p_{\alpha\beta} + p_{\alpha\gamma})}{2} - C_\alpha. \quad (2)$$

The fitness of α , should it not intervene, can be defined as follows:

$$\frac{[p_{\beta\gamma}(p_{\alpha\beta}B_\alpha - Wp_{\alpha\beta}B_\alpha) + p_{\gamma\beta}(p_{\alpha\beta}B_\alpha)]}{2} + \frac{[p_{\gamma\beta}(p_{\alpha\gamma}B_\alpha - Wp_{\alpha\gamma}B_\alpha) + p_{\beta\gamma}(p_{\alpha\gamma}B_\alpha)]}{2}. \quad (3)$$

The term in the first set of squared brackets represents α 's interactions with β . The right-hand term in the numerator is the probability of defeating β , weighted by the appropriate pay-off, given that α did not intervene in a fight between β and γ , and γ went on to win such a fight. The left-hand side of the numerator is the same case, given that β went on to win such a fight (i.e. the fight that α did not break up). The term in the second set of squared brackets represents α 's interactions with γ and should be interpreted in a manner similar to that just described.

Rearranging and simplifying, expression (3) can be rewritten as:

$$\frac{B_\alpha[p_{\alpha\beta}(1 - Wp_{\beta\gamma}) + p_{\alpha\gamma}(1 - Wp_{\gamma\beta})]}{2}. \quad (4)$$

α individuals should then intervene whenever expression (2) > expression (4). This occurs when:

$$\frac{B_\alpha}{C_\alpha} > \frac{2}{(p_{\alpha\beta}p_{\beta\gamma} + p_{\alpha\gamma}p_{\gamma\beta})W}. \quad (5)$$

Inequality (5) demonstrates that increasing winner effects or the strength of α 's position in the hierarchy (reflected by $p_{\alpha\beta}$ and $p_{\alpha\gamma}$) makes intervention by α more likely, whereas increasing the costs of intervention does the opposite.

(ii) Case II: when should β individuals intervene?

Using the same procedure as in expressions (1)–(4), it can be shown that β should intervene whenever the fitness increment accrued for intervention,

$$\frac{B_\beta(p_{\beta\alpha} + p_{\beta\gamma})}{2} - C_\beta, \quad (6)$$

outweighs the fitness increment accrued for not intervening,

$$\frac{B_\beta[p_{\beta\alpha}(1 - Wp_{\alpha\gamma}) + p_{\beta\gamma}(1 - Wp_{\gamma\alpha})]}{2}. \quad (7)$$

Hence, β should intervene in fights whenever

$$\frac{B_\beta}{C_\beta} > \frac{2}{(p_{\alpha\gamma}p_{\beta\alpha} + p_{\gamma\alpha}p_{\beta\gamma})W}. \quad (8)$$

As was the case for α , equation (8) shows that increasing winner effects or the strength of β 's position in the hierarchy (reflected by $p_{\beta\alpha}$ and $p_{\beta\gamma}$) makes intervention by β more likely, whereas increasing the costs of intervention has the converse effect. Moreover, one can assess the relative likelihood of β 's intervention compared with α 's (i.e. is β more or less likely to intervene in fights than α ?). A comparison of the right-hand terms in inequalities (5) and (8) show that the denominator of (8) is less than the

denominator of (5), thus demonstrating that α is more likely to intervene in fights than β , as the B/C ratio for intervention is lower for α individuals.

(iii) *Case III: when should γ individuals intervene?*

γ 's fitness increment for intervening is equal to

$$\frac{B_\gamma(p_{\gamma\alpha} + p_{\gamma\beta})}{2} - C_\gamma. \quad (9)$$

γ 's fitness increment for not intervening can be written as:

$$\frac{B_\gamma[p_{\gamma\alpha}(1 - Wp_{\alpha\beta}) + p_{\gamma\beta}(1 - Wp_{\beta\alpha})]}{2}. \quad (10)$$

γ should then intervene when expression (9) is greater than expression (10), which is equivalent to

$$\frac{B_\gamma}{C_\gamma} > \frac{2}{(p_{\gamma\alpha}p_{\alpha\beta} + p_{\gamma\beta}p_{\beta\alpha})W}. \quad (11)$$

As was the case for α and β , increasing winner effects or the strength of γ 's position in the hierarchy (reflected by $p_{\gamma\alpha}$ and $p_{\gamma\beta}$) makes intervention by γ more likely, whereas increasing the costs of intervention has the converse effect. A comparison of the right-hand terms in inequalities (8) and (11) shows that the denominator of inequality (11) is less than the denominator of (8), thus demonstrating that β is more likely to intervene in fights than is γ .

(b) **Model II: loser effects only**

(i) *Case I: when should α individuals intervene?*

If α intervenes, its fitness increment is the same as in expression (2). If α does not intervene, then its fitness increment equals

$$\frac{[p_{\beta\gamma}(p_{\alpha\beta}B_\alpha) + p_{\gamma\beta}(p_{\alpha\beta}B_\alpha + Lp_{\alpha\beta}B_\alpha)]}{2} + \frac{[p_{\gamma\beta}(p_{\alpha\gamma}B_\alpha) + p_{\beta\gamma}(p_{\alpha\gamma}B_\alpha + Lp_{\alpha\gamma}B_\alpha)]}{2}. \quad (12)$$

The term in the first set of squared brackets represents α 's interactions with β . The right-hand term in the numerator is the probability of defeating β , weighted by the appropriate pay-off, given that α did not intervene in a fight between β and γ , and β went on to lose such a fight. The left-hand side of the numerator is the same case, given that β went on to win such a fight (i.e. the fight that α did not break up). The term in the second set of squared brackets represents α 's interactions with γ and should be interpreted in a similar manner to that just described.

Expression (12) can be rewritten as

$$\frac{B_\alpha[p_{\alpha\beta}(1 + Lp_{\gamma\beta}) + p_{\alpha\gamma}(1 + Lp_{\beta\gamma})]}{2}. \quad (13)$$

It can be shown that expression (2) is greater than expression (13) when

$$\frac{B_\alpha}{C_\alpha} < \frac{2}{(-p_{\alpha\beta}p_{\gamma\beta} - p_{\alpha\gamma}p_{\beta\gamma})L}. \quad (14)$$

Because the right-hand side of inequality (14) is always negative, while the left is always positive, α individuals

should not intervene when loser effects alone are at play in a system.

(ii) *Case II: when should β individuals intervene?*

β 's fitness increment for intervening is defined by expression (6). Its fitness for not intervening is equal to

$$\frac{B_\beta[p_{\beta\alpha}(1 + Lp_{\gamma\alpha}) + p_{\beta\gamma}(1 + Lp_{\alpha\gamma})]}{2}. \quad (15)$$

Expression (6) is greater than expression (15) when

$$\frac{B_\beta}{C_\beta} < \frac{2}{(-p_{\beta\gamma}p_{\alpha\gamma} - p_{\beta\alpha}p_{\gamma\alpha})L}. \quad (16)$$

As in inequality (14), the left-hand part of this inequality is always positive and the right-hand side is always negative. β individuals then should not intervene when loser effects alone operate.

(iii) *Case III: when should γ individuals intervene?*

γ individuals should intervene when the following condition is met

$$\frac{B_\gamma(p_{\gamma\alpha} + p_{\gamma\beta})}{2} - C_\gamma > \frac{B_\gamma[p_{\gamma\alpha}(1 + Lp_{\beta\alpha}) + p_{\gamma\beta}(1 + Lp_{\alpha\beta})]}{2}, \quad (17)$$

which is equivalent to the case where

$$\frac{B_\gamma}{C_\gamma} < \frac{2}{(-p_{\beta\alpha}p_{\gamma\alpha} - p_{\gamma\beta}p_{\alpha\beta})L}. \quad (18)$$

As was the case for α and β , γ individuals then should not intervene when loser effects alone operate.

(c) **Model III: concurrent winner and loser effects**

(i) *Case I: when should α individuals intervene?*

Again, α 's fitness increment, should it intervene, is defined by expression (2). When α does not intervene, its fitness increment when both winner and loser effects are in play is

$$\frac{[p_{\beta\gamma}(p_{\alpha\beta}B_\alpha - Wp_{\alpha\beta}B_\alpha) + p_{\gamma\beta}(p_{\alpha\beta}B_\alpha + Lp_{\alpha\beta}B_\alpha)]}{2} + \frac{[p_{\gamma\beta}(p_{\alpha\gamma}B_\alpha - Wp_{\alpha\gamma}B_\alpha) + p_{\beta\gamma}(p_{\alpha\gamma}B_\alpha + Lp_{\alpha\gamma}B_\alpha)]}{2}, \quad (19)$$

which can be rearranged to

$$\frac{B_\alpha[p_{\alpha\beta}(Wp_{\beta\gamma}(Wp_{\beta\gamma} - Lp_{\gamma\beta}) + p_{\alpha\gamma}(Wp_{\gamma\beta} - Lp_{\beta\gamma}))]}{2}. \quad (20)$$

Then, α should intervene when expression (2) is greater than expression (20), which is true when

$$\frac{B_\alpha}{C_\alpha} > \frac{2}{W(p_{\alpha\beta}p_{\beta\gamma} + p_{\alpha\gamma}p_{\gamma\beta}) - L(p_{\alpha\beta}p_{\gamma\beta} + p_{\alpha\gamma}p_{\beta\gamma})}. \quad (21)$$

(ii) *Case II: when should β individuals intervene?*

β 's fitness increment, should it intervene, is still defined by equation (6). If β does not intervene, its fitness increment is then

$$\frac{[p_{\alpha\gamma}(p_{\beta\alpha}B_{\beta} - Wp_{\beta\alpha}B_{\beta}) + p_{\gamma\alpha}(p_{\beta\alpha}B_{\beta} + Lp_{\beta\alpha}B_{\beta})]}{2} + \frac{[p_{\gamma\alpha}(p_{\beta\gamma}B_{\beta} - Wp_{\beta\gamma}B_{\beta}) + p_{\alpha\gamma}(p_{\beta\gamma}B_{\beta} + Lp_{\beta\gamma}B_{\beta})]}{2}. \quad (22)$$

Expression (22) can be rearranged to

$$\frac{B_{\beta}[p_{\beta\alpha}(Wp_{\alpha\gamma} - Lp_{\gamma\alpha}) + p_{\beta\gamma}(Wp_{\gamma\alpha} - Lp_{\alpha\gamma})]}{2}. \quad (23)$$

Then, β should intervene when expression (6) is greater than expression (23), which is true when

$$\frac{B_{\beta}}{C_{\beta}} > \frac{2}{W(p_{\alpha\gamma}p_{\beta\alpha} + p_{\beta\gamma}p_{\gamma\alpha}) - L(p_{\beta\alpha}p_{\gamma\alpha} + p_{\beta\gamma}p_{\alpha\gamma})}. \quad (24)$$

(iii) *Case III: when should γ individuals intervene?*

Using the same procedure as in expressions (19)–(24), it can be shown that γ should intervene when

$$\frac{B_{\gamma}}{C_{\gamma}} > \frac{2}{W(p_{\gamma\alpha}p_{\alpha\beta} + p_{\gamma\beta}p_{\beta\alpha}) - L(p_{\beta\alpha}p_{\gamma\alpha} + p_{\gamma\beta}p_{\alpha\beta})}. \quad (25)$$

Equations (21), (24) and (25) demonstrate that when winner and loser effects both exist in a system, then it is these values (i.e. W and L) relative to one another that determine whether intervention is favoured. The greater the value of W when compared with L , the greater the probability that intervention will be favoured. As for the case of winner effects alone, absolute intervention probabilities can be ranked as $\alpha > \beta > \gamma$.

3. DISCUSSION

A simple model for the evolution of intervention behaviour as a function of winner and loser effects shows the following three general results:

1. When winner effects alone exist, intervention is favoured by increasing such winner effects, as well as increasing an individual's standing in a hierarchy and/or the benefits of intervention. Naturally, increasing the costs of intervention makes such behaviour less likely. The probability of intervention based on rank is shown to be $\alpha > \beta > \gamma$ (i.e. top-ranked individual > middle-ranked individual > lowest-ranked individual).
2. When loser effects alone are at play in a system, intervention is never favoured, regardless of rank.
3. If both winner and loser effects are operating simultaneously, then the likelihood of intervention behaviour evolving is a function of the relative strength of these two effects. The greater the winner effect relative to the loser effect, the more likely intervention behaviour is to evolve.

When winner effects alone operate, intervention is likely to be favoured because it inhibits others in a group from 'getting on a roll'. That is, intervention by individual i ends encounters that might allow others in the group to emerge victorious during an aggressive interaction, and in the long run to be more of a threat to individual i as a result. Conversely, when loser effects are in play, interventions would break up interactions that could decrease another group member's long-term chances of defeating

any other group member and are therefore selected against. It is worth noting that behavioural endocrinology work has shown that 'aggressive' and 'submissive' behaviours are often controlled by different hormones (Nelson 1995), and thus it is not unreasonable to model winner effects in the absence of loser effects and vice versa. Even if both operate in a given system, however, model III shows clearly that intervention can evolve in such systems, given that the winner effect is relatively potent.

Only recently have behavioural ecologists studied intervention behaviour in animals, and the vast majority of this work has been undertaken in primates (Harcourt & de Waal 1992). By contrast, much of the work on winner and loser effects has not been in primates, but rather, early on in rodents (Ginsburg & Allee 1942; Seward 1946), and more recently in fish (see Chase *et al.* (1994) for a review). Hopefully, the model presented here, which demonstrates how winner and loser effects can drive the evolution of intervention behaviour, will spur future work looking at all these factors simultaneously in primates, fish and many other taxa. Tests of the models presented here could include examining whether populations with pronounced winner versus loser effects contain individuals that intervene in fights more often. One could also work in the opposite direction and study populations that are known to have individuals that intervene and test whether winner effects are stronger than loser effects in such populations.

REFERENCES

- Caro, T. 1994 *Cheetahs of the Serengeti plains*. University of Chicago Press.
- Chase, I., Bartolomeo, C. & Dugatkin, L. A. 1994 Aggressive interactions and inter-contest interval: how long do winners keep winning? *Anim. Behav.* **48**, 393–400.
- Chapais, B. 1995 Alliances as a means of competition in primates: evolutionary, developmental and cognitive aspects. *Yb Phys. Anthropol.* **38**, 115–136.
- Connor, R. 1992 Dolphin alliances and coalitions. In *Coalitions and alliances in humans and other animals* (ed. A. H. Harcourt & F. B. M. de Waal), pp. 415–443. Oxford University Press.
- Connor, R. C., Smolker, R. A. & Richards, A. F. 1992 Two levels of alliance formation among male bottlenose dolphins. *Proc. Natn. Acad. Sci. USA* **89**, 987–990.
- Dugatkin, L. A. 1997 Winner effects, loser effects and the structure of dominance hierarchies. *Behav. Ecol.* **8**, 583–587.
- Ehardt, C. & Bernstein, I. 1992 Conflict intervention by adult male macaques: structural and functional aspects. In *Coalitions and alliances in humans and other animals* (ed. A. H. Harcourt & F. B. M. de Waal), pp. 83–111. Oxford University Press.
- Fentress, J. C., Ryon, J., McLeod, P. J. & Havkin, G. Z. 1986 A multidimensional approach to agonistic behavior in wolves. In *Man and wolf: advances, issues and problems in captive wolf research* (ed. H. Frank), pp. 253–273. Dordrecht: Junk.
- Ginsburg, B. & Allee, W. C. 1942 Some effects of conditioning on social dominance and subordination in inbred strains of mice. *Physiol. Zool.* **15**, 485–506.
- Harcourt, A. H. 1992 Coalitions and alliances: are primates more complex than non-primates? In *Coalitions and alliances in humans and other animals* (ed. A. H. Harcourt & F. B. M. de Waal), pp. 445–472. Oxford University Press.
- Harcourt, A. H. & de Waal, F. B. M. (eds) 1992 *Coalitions and alliances in humans and other animals*. Oxford University Press.
- Landau, H. G. 1951a On dominance relations and the structure of animal societies. I. Effects of inherent characteristics. *Bull. Math. Biophys.* **13**, 1–19.

- Landau, H. G. 1951*b* On dominance relations and the structure of animal societies. II. Some effects of possible social causes. *Bull. Math. Biophys.* **13**, 245–262.
- Maynard Smith, J. & Parker, G. 1976 The logic of asymmetric contests. *Anim. Behav.* **24**, 159–175.
- Nelissen, M. 1985 Structure of the dominance hierarchy and dominance determining ‘group factors’ in *Melanochromis auratus*. *Behaviour* **94**, 85–107.
- Nelson, R. 1995 *An introduction to behavioral endocrinology*. New York: Sinauer Press.
- Packer, C. & Pusey, A. E. 1982 Cooperation and competition within coalitions of male lions: kin selection or game theory? *Nature* **296**, 740–742.
- Parker, G. & Rubinstein, D. 1981 Role assessment, reserve strategy and the acquisition of information in asymmetric animal contests. *Anim. Behav.* **29**, 221–240.
- Parker, G. A. 1974 Assessment strategy and the evolution of fighting behaviour. *J. Theor. Biol.* **47**, 223–243.
- Pusey, A. E. & Packer, C. 1995 The ecology of relationships. In *Behavioural ecology*, 4th edn (ed. J. Krebs & N. B. Davies), pp. 254–283. Oxford: Blackwell Scientific.
- Russell, J. K. 1983 Altruism in coati bands: nepotism or reciprocity? In *Social behavior of female vertebrates* (ed. S. Wasser), pp. 263–90. New York: Academic Press.
- Seward, J. P. 1946 Aggressive behavior in the rat. IV. Submission as determined by conditioning, extinction and disuse. *J. Comp. Psychol.* **39**, 51–57.
- Silk, J. B. 1992 Patterns of intervention in agonistic contests among male bonnet macaques. In *Coalitions and alliances in humans and other animals* (ed. A. Harcourt & F. B. M. de Waal), pp. 214–232. Oxford University Press.
- Zabel, C., Glickman, S., Frank, L., Woodmansee, K. & Keppel, G. 1992 Coalition formation in a colony of prepubertal spotted hyenas. In *Coalitions and alliances in humans and other animals* (ed. A. H. Harcourt & F. B. M. de Waal), pp. 112–135. Oxford University Press.