Animal Behaviour 82 (2011) 295–302

Contents lists available at ScienceDirect

Animal Behaviour

journal homepage: www.elsevier.com/locate/anbehav

Long-term commitment promotes honest status signalling

Szabolcs Számadó*

HAS-ELTE Research Group for Theoretical Biology and Ecology, Department of Plant Taxonomy and Ecology, Eötvös Loránd University

ARTICLE INFO

Article history:
Received 12 May 2010
Initial acceptance 9 September 2010
Final acceptance 20 April 2011
Available online 8 June 2011
MS. number: 10-00335R

Keywords:
badge of dominance
commitment
cost-free signal
honest signalling
status signalling

Although there are many examples of status signalling in nature, the mechanisms that maintain the stability of these systems are still poorly understood. Here I show in terms of a simple game of aggressive communication that commitment to the need to defend a given resource repeatedly in the long term greatly increases the conditions under which signalling of strength by means of conventional badges can be honest and evolutionarily stable. Such commitment can efficiently prevent the invasion of potentially weak cheaters pretending to be strong at a very low cost of harassment suffered by such weak individuals. Moreover, as long as such commitment is present, this cost is not a function of the contested resource; thus the value of the resource can be orders of magnitude higher than the potential cost imposed on cheaters, yet the system remains honest and evolutionarily stable. While liar–strong strategies (i.e. Trojans) can invade under some conditions, in most cases there is a broad region of honesty and this honest region is always open ended towards large resource values. In this region of honesty, no other costs, either production costs or handicaps, are necessary to maintain the evolutionary stability of the system.

The honesty of badges of dominance is a long-standing issue. Badges can signal the dominance of a bird within a flock (Rohwer & Rohwer 1978; Møller 1987; Liker & Barta 2001) or the strength of the owner of a territory (Part & Qvarnström 1997; Garamszegi et al. 2006; reviewed in Senar 1999). Despite the obvious temptation to cheat, when weak or submissive individuals display large badges to claim the resource without fighting, badges are found to be reliable both in birds (Rohwer & Rohwer 1978; Møller 1987; Senar 1999) and in other species such as paper wasps, *Polistes dominulus* (Tibbetts & Dale 2004; Tibbetts & Lindsay 2008; but see Cervo et al. 2008) and sand lizards, *Lacerta agilis* (Olsson 1994). The source of this reliability is the topic of debate here. Two main alternative proposals can be found in the literature (reviewed in Senar 1999). The 'social control' hypothesis predicts that the high cost of fighting among dominants prevents cheating. This cost can be high either because there are frequent fights among dominants or because fighting among dominants is particularly costly even if it is not more frequent than expected on a random basis. Some authors have indeed found that dominants fight more frequently with each other (Jarvi & Bakken 1984; Møller 1987); others, however, have found the opposite to be the case (Keys & Rothstein 1991; Slotow et al. 1993). While fighting between dominants was found to be costly in some species (Keys & Rothstein 1991) this does not seem to be a general pattern (Senar 1999). Although the original 'social control' hypothesis found support from Rohwer (1977), later Rohwer himself proposed an alternative solution. According to the 'incongruence' hypothesis (Rohwer & Rohwer 1978) birds increase their level of aggression if they observe incongruent signals.

Since none of the previous proposals seemed to give a satisfactory solution, other costs unrelated to aggressive communication were proposed including increased resting metabolic rate for dominants (Jarvi & Bakken 1984), avoidance of dominant individuals by subordinates (Senar & Camerino 1998), higher predation risk for dominants because of their more conspicuous colours (Keys & Rothstein 1991; Slotow et al. 1993) and immunocompetence handicaps (i.e. cost of melanin displays imposed on honest signallers mediated by the immunosuppressive effect of testosterone, Folstad & Karter 1992). However, these costs received mixed empirical support as well (Senar 1999; Roberts et al. 2004).

Theoretical models investigating the issue made the next two relevant predictions. (1) The badges of dominance are expected to evolve only if the cost of fighting is greater than the value of the contested resource (Maynard Smith & Harper 1988). (2) The so-called 'Trojan' strategy (where strong individuals pretend to be weak) can always invade unless there is a contest-independent cost to dominance (Owens & Hartley 1991; Johnstone & Norris 1993).

Independently of these models, Enquist (1985) demonstrated in a simple game of aggressive communication that honest signalling of strength by means of cost-free conventional signals can be evolutionarily stable. Enquist's model never featured prominently

* Correspondence: S. Számadó, HAS-ELTE Research Group for Theoretical Biology and Ecology, Department of Plant Taxonomy and Ecology, Eötvös Loránd University, Pázmány Péter sétány 1/c, Budapest H-1117, Hungary.
E-mail address: szamszab@ludens.elte.hu.

0003-3472/$38.00 © 2011 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.anbehav.2011.04.024

in the debates on the honesty of badges of dominance; nevertheless, it is the simplest model of aggressive communication that allows honest conventional signalling (Hamblin & Hurd 2007). The key assumption of the model is that it requires commitment to the signal, that is, it assumes that those signallers that give signals indicative of high strength cannot get away without fighting when the enemy attacks even if they want to do so. In other words, weak cheaters signalling strength have to pay the cost of fighting even if they would rather escape. How this commitment is generated is the key issue of honesty in this model. For example, Számadó (2008) was able to show that proximity to the rival during threat displays can generate this commitment. Here, by walking into the range of the 'honest striking distance' (Számadó 2008) of the opponent, the individual voluntarily gives up the option to flee without fighting. This changes the cost structure of the game, which in turn promotes honesty. Dominance displays, however, clearly work differently from threat displays (Walther 1984). Namely, dominance displays are one-to-many displays as opposed to the one-to-one nature of threat displays, and accordingly they can be given from much larger distances, so much so that the opponent might be (although need not be) completely out of sight (i.e. birdsong or rutting of deer). This implies, in turn, that proximity risk during dominance signalling cannot always create the commitment necessary for Enquist's model to work.

Another problem, and not just with Enquist's model but with all the previous models in general, is that they investigate single-shot interactions. Dominance relations, however, usually involve repeated interactions in which an individual has to win a large proportion of the fights to enjoy the benefits of dominance (i.e. it has to establish itself as a dominant, see dominance matrices in Fugle et al. 1984; Rohwer 1985; Watt 1986; Møller 1987; Slotow et al. 1993). The same observations concerning repeated interactions and the need to win a number of fights hold for other contexts as well where badges of dominance systems have been described, such as mate or territorial defence (Part & Qvarnström 1997).

Here I model the effect of commitment generated by repeated interactions on the evolution of signal honesty and investigate how this effect will change the output of the model with regard to the two previously described predictions of the existing models. This commitment can be made truly 'binding' in the case of birds by the fact that they moult their badge once a year; hence they have to carry on with the signal they 'chose' at the time of moulting even if later on they would prefer to change it (although this kind of 'binding' is not necessary for the model to work). The model is introduced in the next section. Based on the model I investigate first the stability of the honest strategy against weak cheaters signalling strength (i.e. against traditional 'cheaters'), then against strong cheaters signalling weakness (i.e. against 'Trojans').

THE MODEL

Here I use Enquist's (1985) model but with two crucial differences. First, individuals engage in repeated interactions, they fight n fights, and they have to be able to win above a certain threshold (k) to gain possession of the resource (where $k = < n$). Second, since we model dominance displays that are usually given from much larger distances than threat displays, we assume that individuals can flee even when the opponent attacks, incurring only a minimal cost (F_f). That is, unlike in Enquist's original model, individuals that are attacked do not have to pay the full cost of fighting. Enquist's model can be seen as a modified version of the Hawk–Dove game (Maynard Smith 1982), in which each player can be weak or strong and knows its own strength but not that of the opponent. The game has two steps. In the first step each player can choose between two cost-free signals A or B; then in the second round each animal can give up, attack unconditionally or attack if the opponent does not withdraw. Let V denote the value of the contested resource, and C_{WW} and C_{SS} the expected costs of a fight between two weak and two strong individuals, respectively. We assume that a strong animal can always beat a weak one with a cost C_{SW}, and C_{WS} is the expected cost that a weak animal should suffer on this occasion. The following relation holds between these costs: $C_{WS} > C_{WW}$, $C_{SS} > C_{SW}$. Let us denote the cost of fleeing as F_f and the cost of attacking a fleeing opponent as F_A, and further suppose that there is a cost of waiting if the opponent attacks unconditionally (F_P). It is biologically realistic but not necessary to assume that $C_{SW} > F_A$ and $C_{SW} > F_P$ (Hurd 1997); let us assume in line with our original argument that the cost of fleeing is small; thus $C_{WS} > F_f$. Individuals of equal strength have an equal chance of winning against each other. The frequencies of weak and strong individuals are denoted by q and $1 - q$, respectively. Then the payoffs for weak and for strong contestants can be written as shown in Table 1.

CHEATERS AND COMMITMENT

Enquist (1985) was able to show that the following global strategy (S_a) is evolutionary stable.

Strong individuals should show A in the first round and then attack unconditionally if the opponent shows A or wait until the opponent flees if it has shown B. Weak individuals should signal B at the first step and then attack unconditionally if the opponent shows B or withdraw if the opponent signals A.

Strategy S_a is a pure strategy in which both strong and weak animals signal honestly. The corresponding dishonest pure strategy (S_b) can be defined as follows. Always display A in the first round, regardless of strength; then in the second round, if strong attack unconditionally if the opponent shows A or wait until the opponent

Table 1
The expected payoffs of the strategies in a model of aggressive communication

Ego strength		Opponent strength					
		Strong			Weak		
		Attack	Conditional attack	Flee	Attack	Conditional attack	Flee
Strong	Attack	$0.5V-C_{SS}$	$0.5V-C_{SS}$	$V-F_A$	$V-C_{SW}$	$V-C_{SW}$	$V-F_A$
	Conditional attack	$0.5V-C_{SS}-F_P$	$0.5V-C_{SS}$	V	$V-C_{SW}-F_P$	$V-C_{SW}$	V
	Flee	$-F_f$	0	$0.5V-C_{SS}$	$-F_f$	0	$0.5V$
Weak	Attack	$-C_{WS}$	$-C_{WS}$	$V-F_A$	$0.5V-C_{WW}$	$0.5V-C_{WW}$	$V-F_A$
	Conditional attack	$-C_{WS}-F_P$	$-C_{WS}$	V	$0.5V-C_{WW}-F_P$	$0.5V-C_{WW}$	V
	Flee	$-F_f$	0	$0.5V$	$-F_f$	0	$0.5V-C_{WW}$

Modified after Számadó (2000). Each cell gives the fitness of the Ego player playing against the strategy played by the column player. V: value of the contested resource; C_{SS}, C_{WW}: expected costs of fight between equal opponents; C_{SW}: cost for strong individual to beat weak one; C_{WS}: cost to weak individual when beaten by strong one; F_f: cost of fleeing; F_A: cost of attacking fleeing opponent; F_P: cost of waiting if opponent attacks unconditionally.