ELSEVIER
ANIMAL BEHAVIOUR, 2005, **70**, 1155–1170
doi:10.1016/j.anbehav.2005.02.014

A strategic taxonomy of biological communication

PETER L. HURD* & MAGNUS ENQUIST†
*Department of Psychology, University of Alberta
†Department of Zoology, University of Stockholm

*(Received 6 March 2003; initial acceptance 10 June 2003;
final acceptance 17 February 2005; published online 3 October 2005; MS. number: A9569R)*

The rapid increase in the understanding of biological communication has been largely theory driven. Game-theoretical models have completely changed how behavioural biologists think about communication. These theories are widely cited, but much of the theoretical work remains only vaguely understood by the wide majority of those who make use of it. Critical terms and concepts have not been sufficiently well defined to provide a solid common foundation for theoretical work as a unified whole. In this paper, we synthesize the relevant theoretical work, and describe its relation to more classical thinking within the study of animal behaviour. We present a taxonomy of biological communication and signals. This taxonomy is based upon the structure and properties of extensive-form game models of communication. We then classify many influential models of communication according to the structure of their extensive-form representations. Finally, we link signal taxonomy and extensive form to show how the temporal structure of a modelled interaction determines the type of signal the model will produce. We argue that this intuitive form of modelling is the correct one for unifying terminology and theory.

Recent years have seen a dramatic increase in understanding of biological communication, one of the longest studied aspects of animal behaviour. These advances are largely the result of a number of game-theoretical models, which have provided a clear understanding of the issues at the core of communication (Maynard Smith 1974; Maynard Smith & Parker 1976; Caryl 1979; Enquist 1985; Maynard Smith & Harper 1988; Grafen 1990). This modelling tradition, the strategic perspective, is the foundation for many present views on the evolution of communication (Bradbury & Vehrencamp 1998; Johnstone 1998; Maynard Smith & Harper 2004). In this paper, we synthesize these theoretical advances. In particular, we propose a taxonomy of biological communication based on strategic analyses of signals and signalling interactions.

How animals use and respond to signals are traits whose evolution is frequency dependent (Maynard Smith 1982). An individual's success or failure in using signals depends upon how other individuals use and interpret signals. It is commonly assumed that traits under frequency-dependent selection will evolve to an evolutionarily stable endpoint. The aim of evolutionary game theory is to predict and describe such stable equilibria (Maynard Smith 1982; Houston & McNamara 1999). Most signalling games consider pairwise interactions between individuals chosen at random from the population, and we limit our description of evolutionary stability to this situation.

Let $w(s',s'')$ denote the fitness of using strategy s' against an individual who uses strategy s''. Consider a population of individuals that uses a strategy s^* and consider also a rare mutant strategy s'. If s' has higher fitness than s^*, the mutant can invade the population. Thus, a necessary condition for evolutionary stability is that

$$w\left(s^*,s^*\right) \geq w\left(s\prime,s^*\right) \text{ for all } w\prime \tag{1}$$

This is referred to as a Nash equilibrium (Fudenberg & Tirole 1991). One way of expressing this is that s^* must be a strategy that is the best strategy to use or play against itself. An individual in a population of s^* strategists cannot benefit by choosing any alternative to s^*. Two additional requirements are frequently applied to Nash equilibria to ensure that they are evolutionarily stable (Houston & McNamara 1999). The first is Maynard Smith's (1982) concept of evolutionary stability. Evolutionarily stable strategy (ESS) conditions ensure stability against invasion when mutants are equally fit (condition 2, below). According to Maynard Smith, s^* is an ESS if for all $s' \neq s^*$, one of the following two conditions holds:

$$\text{(condition 1)}\quad w\left(s^*,s^*\right) > w\left(s\prime,s^*\right) \tag{2}$$

Correspondence: P. L. Hurd, Department of Psychology, University of Alberta, AB T6G 2E9, Canada (email: phurd@ualberta.ca). M. Enquist is at the Department of Zoology, University of Stockholm, Stockholm, S-106 91, Sweden.

0003–3472/05/$30.00/0

$$\text{(condition 2)}\quad w\left(s^{*},s^{*}\right)=w\left(s\prime,s^{*}\right)\quad\text{and}\quad w\left(s^{*},s\prime\right)>w(s\prime,s\prime). \tag{3}$$

The second stability requirement is convergence stability (Eshel 1983). Convergence requires that a population returns to s^* if slightly perturbed from this state (Houston & McNamara 1999). From this second requirement it also follows that evolution can lead to s^*. Maynard Smith's ESS condition is sufficient to guarantee convergence when the number of possible strategies is finite. However, evolutionary stability requires that additional convergence conditions be met when strategy variables are continuous (Eshel 1983).

Signalling games are often formulated as interactions between two types of individuals, senders and receivers. Games between individuals with differing roles are known as asymmetric games (Selten 1980; Maynard Smith 1982; Leimar & Enquist 1984). In analysing such games one may seek a set of strategies, one per role, rather than a single stable strategy conditioned on role. The requirements for evolutionary stability are essentially the same as above. Sender and receiver strategies must be the best ones to use against each other, and this pair must be stable against invasion and perturbations.

Developments in evolutionary modelling beyond the scope of this paper (see e.g. Eshel 1996; Geritz et al. 1998; Houston & McNamara 1999) have increased and modified our understanding of evolutionary stability. More importantly, many important evolutionary problems lack stable equilibria. Dynamic approaches, which embrace nonequilibrium solutions, are required to understand the evolution of the phenotypes under such circumstances. So far, however, these developments have had only minor effect on game-theoretical modelling of biological signalling. That signalling strategies may remain in constant evolutionary flux is a real possibility, and we consider such signals an important part of our taxonomy.

This strategic analysis of signals and communication systems is based upon the extensive-form game representation of interactions (Fudenberg & Tirole 1991; Gibbons 1992). Extensive-form game theory is so useful in modelling communication because its structure (i.e. sequences of behavioural decisions) matches the time flow of interactions between individuals. This correspondence provides a natural framework for both classifying the type of interaction in which communication takes place and characterizing the nature of the signals that are used. Addressing the extensive form of biological interactions also allows for the classification of signal properties in a manner that is relevant to understanding their evolution. The extensive form also lets us define previously slippery hermeneutic concepts such as 'message' and 'meaning' (Smith 1977) in terms of, respectively, signal and receiver preference ranking (Trivers 1974; Maynard Smith 1991) and as a function of hidden state. (These and other terms are defined in the Appendix.)

A simple example of an extensive-form mate signalling game between spiders is presented in Fig. 1. Play begins at the open circle, the node at the top of the graph, and flows down to the terminal nodes where payoffs are assigned to each player, w(F) to the female and w(M) to the male. The player making a decision (i.e. the player whose turn it is to move) is noted beside each node. In this game the players are male (M), female (F) and nature (N). Nature is a 'player' whose only move is to determine which state the female is in for any given play for the game. The female is in one of three states: receptive (R), hungry (H) or disinterested (D), and signals to the male with one of two signals, 'go away' (G) or 'come here' (C). The probabilities of receptive, hungry and disinterested moves by nature are x, y and $1 - (x + y)$, respectively. These probabilities are parameters, fixed and known to both players. Dotted lines join information sets (i.e. nodes that the current player is not capable of distinguishing between). In this case, the male cannot tell what state the female is in, only which signal she has chosen. The male responds to the female's signalling move by either approaching or fleeing (A or F). The male's choice of behaviour must be the same for all nodes within an information set. The heart of the game is how the receiver's beliefs about the probable signaller state changes given the signal. A different set of

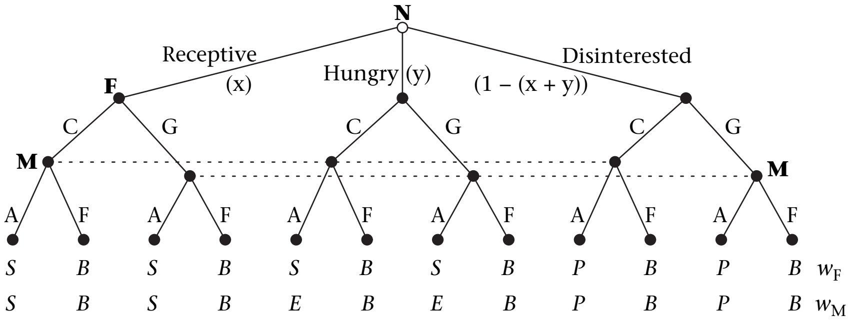


Figure 1. Extensive-form mate signalling game between spiders. The players are the female (**F**), the male (**M**) and nature (**N**). A female spider is either receptive, hungry or disinterested with Pr(x, y, $1 - (x + y)$). She signals with one of 'come here' (C) or 'go away' (G). Her payoffs are satisfied (S), baseline (B) or Pestered (P). The male then either approaches (A) or flees (F), and his payoffs are satisfied, baseline, pestering or eaten (E). For the female payoffs, $S > B > P$, and for the male, $S > B > P > E$. Nodes that cannot be differentiated by the male are joined by a dotted line. The signals in this game are costless; the payoff consequences to the signals in that game are because of the differing responses that they elicit.

beliefs will exist before the signal (which will be the probabilities of the respective moves by nature) and at each of the information sets after a signal is observed. The communication equilibrium solution to this game is discussed later.

The organization of the next section of this paper follows the proposed taxonomy of signal types shown in Fig. 2. This taxonomy includes and builds upon earlier work, most directly that of Maynard Smith & Harper (1995), Hurd (1997a) and Vehrencamp (2000). Each branch has its own section, describing the principal differences between the alternatives as well as providing examples. Table 1 relates each of these theoretically defined signal types to an example and likely empirical properties. The following section describes a classification of models based on their extensive-form structure and the association between signal type and game structure.

Signal taxonomy

The first distinction that our signal taxonomy makes is whether signalling behaviours are evolutionarily stable. Such signals must be understood differently from those that constantly change over evolutionary time. The next distinction is whether the choice of signal is constrained. Strategic signals, those whose use is not constrained and thus with which 'bluffing' is possible, may further be classified according to the mechanism that maintains their evolutionary stability. This mechanism may be due to (1) handicapping signal costs, (2) the socially derived costs of the conventional signals or (3) they may be a hybrid of the two, an interaction handicap.

Stability versus Constant Evolutionary Change

When signalling behaviour is evolutionarily stable, as defined in the introduction, no player can improve on its expected outcomes by using an alternative strategy. Both senders and receivers behave optimally given the other's choice of behaviour.

Stability

An important consequence of stability is that we can interpret communication as a transfer of knowledge. Game-theory definitions of evolutionary stability explicitly require mutual anticipation of strategies between the players at a stable outcome. The rule used by the signaller in choosing a signal must match the receiver's expectations of how the signaller will behave in any possible situation. The receiver must know the rule that the signaller uses in choosing a signal. When there is such mutual understanding between individuals, we can apply

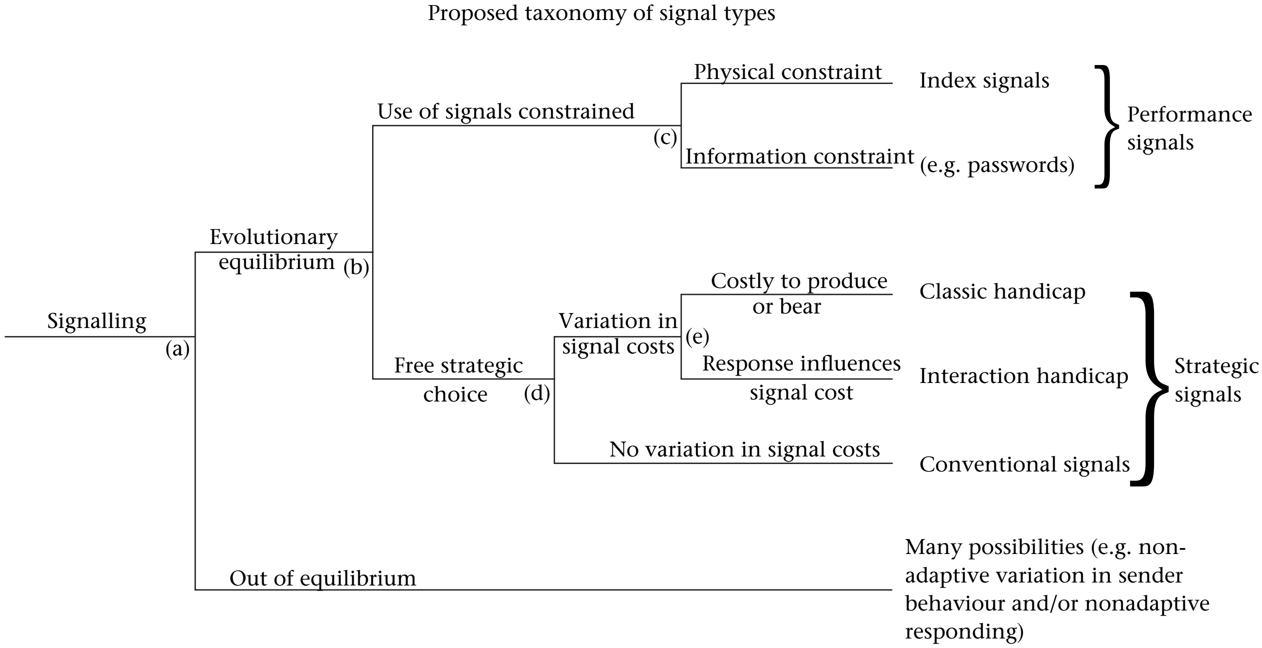


Figure 2. The proposed classification of biological signals based on factors that maintain variation in sender and receiver behaviour. Signalling cannot exist without such variation. This pattern of signal variation and response variation (a) is either at an evolutionary equilibrium (either stable or unstable) or it is out of equilibrium. When signalling is out of evolutionary equilibrium, variability in responses to signals may occur because of biases and variability in signal choice that reveals information to the detriment of the signaller. Signals that are used at an evolutionary equilibrium (b) are either performance signals, whose use is constrained so that only some signallers are able to use some signals, or strategic signals, which are subject to free strategic choice. In the latter condition all signallers may use all signals, and the choice of signal is determined by the relative costs and benefits of each signal given the state of the signaller. Performance signals (c) may be index signals, whose use is constrained by physical ability to perform the signal, or signals whose use is constrained by having access to some information. Signals that are subject to free strategic choice (d) are either handicapped by signal cost (cost of producing or bearing the signal) or conventional signals. Handicapped signals (e) have a cost to their production or in the bearing of the signal. This cost is either paid regardless of how the receiver responds (classic handicaps) or paid as a function of how the receiver responds (interaction handicaps).

Table 1. Signal properties as a function of signal type*

	Signal type			
		ESS		
			Strategic signalling	
Property	Out of equilibrium	Performance signalling	Handicap	Conventional
Transfer of information?	Combination: manipulation and information transfer	Yes	Yes	Yes
Conflict required for category to exist?	Yes	Yes	Yes	No
Form of signal	Exaggerated form	May not look like a signal	Exaggerated form	Unexaggerated
Cost of signal production	Potentially high	No prediction, may be zero	Potentially high	Zero
Examples	'Sword' length (Basolo 1990)	Lateral display (Enquist & Jakobsson 1986)	Peacock tail (Petrie 1994)	Song matching (Molles & Vehrencamp 2001)

*Probable properties of signal form within each major category. Each category in our taxonomy generates definite and often unique predictions about signal properties.

Smith's (1977) hermeneutic-inspired language of 'message' and 'meaning' and credibly suppose that both signaller and receiver associate the same signals to the same messages. For instance a coot, *Fulica americana*, making use of the 'bracing' display (Gullion 1952) is communicating the message 'bracing display' to a conspecific with whom it is presumably engaged in an evolutionarily stable signalling exchange. To a passing great-horned owl, *Bubo virginianus*, with whom it is not engaged in an evolutionarily stable signalling exchange, it is not communicating the message 'bracing display' but is merely signalling its presence.

It is not necessary that the receiver be able to interpret the signaller's state perfectly. For instance, imagine the case of a forager listening for an alarm call from a sentry. The sentry gives an alarm call under two conditions, either when a predator is detected or when it wishes to steal a large prey item from a forager (Munn 1986). An evolutionarily stable communication system may exist in which half the alarm calls are 'false'. The forager may have the prior belief that any given alarm call has only a 50% probability of signalling a predator, yet it may be worth reacting to them all as if they were 100% accurate. If the receiver's beliefs about probable signaller states are accurate, then the evolutionarily stable signal has the meaning 'drop everything and run away', and not 'predator present, run away'.

A set of ESSs (evolutionarily stable strategies) exists for the spider game (Fig. 1), which has this property. We can describe a female's strategy as the probability of using the C signal in states R, H and D, and the male's strategy as the probability of playing A after signals C and G. The strategy profile $\text{Female}_{(1,1,0)}$, $\text{Male}_{(1,0)}$ is that in which the female uses the signal 'come here' when either receptive or hungry and 'go away' when disinterested, and males approach females who signal 'come here' and flee from females who signal 'go away'. These moves lead to outcomes. For males, these outcomes are Satisfied, Baseline, Pestering and Eaten, with corresponding payoffs S, B, P and E. For females, the outcomes are Satisfied, Baseline and Pestered, with corresponding payoffs S, B and P. Players at this profile receive payoffs of

$$F_{\text{Female}(1,1,0)} = xS + yS + \big(1-(x+y)\big)B \tag{4}$$

$$F_{\text{Male}(1,0)} = xS + yE + \big(1-(x+y)\big)B \tag{5}$$

but alternative strategies for males pay

$$F_{\text{Male}(1,1)} = xS + yE + \big(1-(x+y)\big)P \tag{6}$$

$$F_{\text{Male}(0,0)} = xB + yB + \big(1-(x+y)\big)B = B. \tag{7}$$

Males do not benefit from changing strategy as long as $B > P$ and $xS + yE > xB + yB$. It is assumed that $S > B > E$, so the condition is met when y is sufficiently smaller than x.

Alternative strategies for females are not as interesting as for males but include

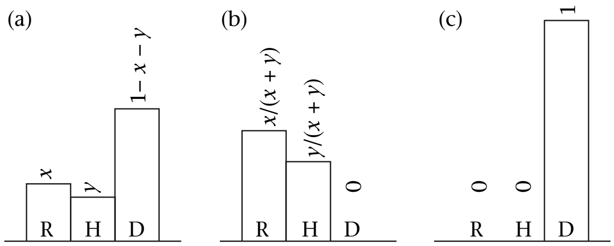
$$F_{\text{Female}(1,1,1)} = xS + yS + \big(1-(x+y)\big)P \tag{8}$$

$$F_{\text{Female}(1,0,0)} = xS + yB + \big(1-(x+y)\big)B \tag{9}$$

$$F_{\text{Female}(0,1,0)} = xB + yS + \big(1-(x+y)\big)B. \tag{10}$$

Females do not benefit from changing to any of these strategies as long as $B > P$ and $S > B$. Other female strategies exist (e.g. $\text{Female}_{(1,0,1)}$), but these can easily be shown to be inferior to $\text{Female}_{(1,1,0)}$ under the same conditions.

We can describe the information transferred by an evolutionarily stable signal by the inferred probability of each female state given the male's location within the extensive-form game tree (Fig. 3). The male has different



Before signal After signal C After signal G

Figure 3. The male's perceived probability that the female spider is either receptive (R), hungry (H) or disinterested (D) as a function of her signal when the strategy profile is ($\text{Female}_{(1,1,0)}$, $\text{Male}_{(1,0)}$). On the left is the belief prior to a signal, in the middle the updated belief after receiving a C signal, and on the right is the updated belief after a G signal.

beliefs about female state following each of her signals. Before the signal, the male has the prior belief that the female is in states R:H:D with probabilities $x{:}y{:}1-(x+y)$; after signal G his belief is 0:0:1 and after signal C his belief is $x/(x+y){:}y/(x+y){:}0$. Note that communication is evolutionarily stable without the condition of perfect information, or 'honesty', in that the receiver does not know whether a C signal indicates state R or H, because these two states pool in using the same signal. The receiver has, however, gained enough information to decide between a G or C response without the signaller using an unambiguous signal. The signal used by females in the H state may be said to be 'dishonest' (sensu Johnstone & Grafen 1993), because the response is consistently to the female's benefit and to the detriment of the male. But this is not to say that the male is necessarily fooled into 'misinterpreting' the signal. If the subsequent belief after a C signal matches the probability states in Fig. 3, then there is an equilibrium interpretation, and the meaning (Smith 1977) of the signal is 'come here' rather than 'receptive'. Suppose that a parent feeds begging nestlings in its nest with the belief that the probability that they are nest parasites is the actual probability of having nest parasites. Then an equilibrium exists, and the meaning of begging is 'feed me', rather than 'I am your offspring; feed me'. For this reason we dislike the terms 'honest' and 'dishonest' and prefer to use 'reliable' and 'unreliable', or 'unambiguous' and 'ambiguous' whenever possible.

Communication under constant evolutionary change

The previous section dealt with communication systems in evolutionary equilibrium. It has long been recognized that biological signalling systems may not be at an evolutionary equilibrium (Fisher 1930; Dawkins & Krebs 1978; Krebs & Dawkins 1984; Enquist et al. 2002). Several cases have been recognized. The first is a persuasion/resistance arms-race, in which senders continuously develop new signals to manipulate receivers, and receivers counteradapt continuously to these changes in signal form (Williams 1966; Dawkins & Krebs 1978; Arak & Enquist 1995). Another case is Batesian mimicry, in which the signaller elicits a favourable response from the receiver by mimicking another type of signaller. In such evolution, the receiver may be continuously evolving improved discrimination, and the mimic sender is selected to avoid being discriminated against (Dittrich et al. 1993). Some models of sexual selection and sexual conflict also suggest that evolution leads to a constant out-of-equilibrium dynamic (e.g. Fisher 1930; Parker 1979; Lande 1981; Chapman et al. 2003). These are but a few well-known examples; there are many other possibilities and many possible nonequilibrium evolutionary dynamics.

In all of these nonequilibrium examples, animals are still communicating; the signals influence the behaviour of the receiver. The major difference between these examples and equilibrium communication systems is that we cannot interpret nonequilibrium situations as a transfer of information in any precise way. In equilibrium conditions, receivers may interpret signals by updating ambiguous beliefs about state to new, but ambiguous, beliefs in a Bayesian manner. The reception of ambiguous but statistically accurate information allows receivers to make better decisions. Out of equilibrium, the receiver's response may not be the most beneficial to itself because the receiver acts as though the signal indicates something it does not. This makes an interpretation of signal 'meaning' (Smith 1977) more problematic, if not impossible.

For example, consider an aposematic species in which all individuals are equally toxic, but which vary in the intensity of their red warning coloration. Assume that the receivers are sensitive to the intensity of the signal and are deterred more often by redder signals. That the signal is red, and not palatable blue, is informative, but the variation in intensity of redness has no real meaning. Redness does not inform about the amount of danger, other than that some red means danger. Many biological signals seem to be overly complex in this manner (e.g. Enquist et al. 2002).

The issues raised by communication out of evolutionary equilibrium are important ones. We will not present any new theoretical treatment of these issues here. Discriminating between systems that are evolutionarily stable and those that are not is not a trivial task. The distinction requires a judgement call based on a sound biological understanding of the system, rather than a simple empirical test.

Performance versus Strategic Signalling

Given that signalling is evolutionarily stable, we must have variation in both signal use and in the response to signals. The subsequent criteria in our signal taxonomy all pertain to how that variation in use is maintained. The maintenance of this variation is the central problem of biological communication. If there is a favoured response from the receiver, why does the signaller not always use whatever signal is most likely to elicit this response?

Performance signals

Maynard Smith (1982) suggested a simple mechanism that may stabilize variation in some threat displays (i.e.

that signals were constrained or 'unbluffable' such that some signallers were able to use signals that other signallers were not). If a signal's performance is constrained to a subset of signallers, we call this a performance signal (Enquist 1985). Not all signallers are capable of producing all signals; some signallers are constrained to choosing between a subset of signals (Fig. 4).

Two general subclasses of performance signals can be identified, those based on physical ability and those based on information. Signals based on physical ability are called 'index signals' (Maynard Smith & Harper 1995; Vehrencamp 2000), and we call signals whose use is constrained by the signaller's access to information 'information-constrained signals'. The extensive-form representation of both these forms of performance signal is the same.

A biological example of index signalling of physical ability is the croaking of some frogs. In many species, the croak of large frogs is deeper than that of small frogs, because the lower frequencies are simply unattainable to smaller frogs (Martin 1972; Gerhardt 1994) and are therefore reliable through a constraint on performance. The lateral display, tail-beat and mouth wrestle behaviours of the cichlid fish *Nannacara anomala* all appear to be signals of this type. A lateral display, in which a fish poses side-on to its opponent, is unbluffable. Although all fish may attempt to look as large as possible, small fish will look smaller than large fish. While tail beating, a fish directs a jet of water towards its opponent (Jakobsson et al. 1979; Enquist et al. 1990), the force of which provides tactile cues as to the fish's weight. A small or weak fish is physically incapable of generating as much force as a large fish. Mouth wrestling fish lock jaws and attempt to push each other backwards (Baerends & Baerends-van Roon 1950). The physical constraint of ability means that smaller fish cannot indicate any strength they choose; they signal their weaker state by pushing with less force than a big fish can.

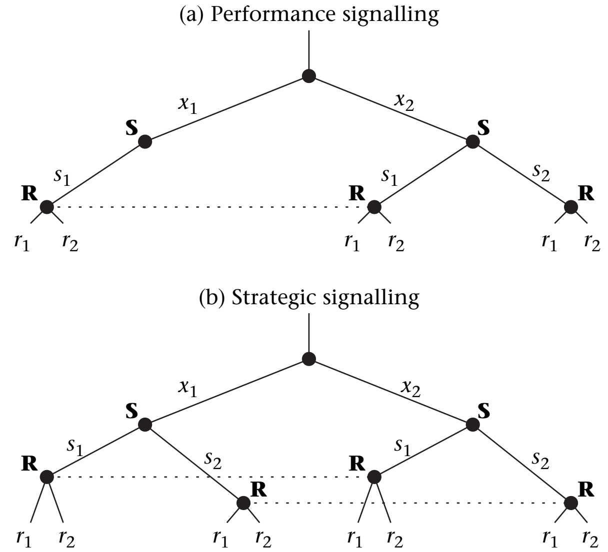


Figure 4. A schematic of performance and strategic signals using extensive-form game trees. In (a) performance signalling, signallers (**S**) in state x_1 are constrained to use signal s_1, but signallers in state x_2 are free to choose between signals s_1 and s_2. A receiver (**R**) knows that the signaller must be in state x_2 whenever s_2 is used, but does not know with certainty whether the signaller is in x_1 or x_2 when s_1 is used. In (b) strategic signalling, all signallers may use all signals. The relative benefits of alternative signals will depend upon the cost/benefit trade-offs of different outcomes.

An example of an information-constrained signal is stalked prey staring at a hidden predator, thereby signalling to the predator both its alerted state and the futility of continuing the hunt. The signal can be performed only by a signaller who knows the location of the hidden predator. Like a password, it is available to potential signallers only if they have some particular information.

Performance signalling has also been called unambiguous signalling (Maynard Smith 1982), unbluffable signalling (Wiley 1983), assessment signalling (Maynard Smith & Harper 1988; Johnstone 1998) and revealing handicaps (Grafen 1990). The semiotic term 'index' certainly predates 'performance signalling' outside the biological literature, and has since become widely accepted within biological communication literature (Maynard Smith & Harper 1995). When index signal has been defined explicitly (e.g. Bradbury & Vehrencamp 1998; Maynard Smith & Harper 2004) it has referred to physical or physiological constraint. We see index signals as a subset of a more general class of performance signals.

Strategic signals

In contrast to performance signals, strategic signals are available to all signallers. Variation in the signal comes from variation in the choice of the signal, rather than the performance of the signal. The choice of an optimal signal is a strategic decision based on the relative costs and benefits of all the available alternatives. Understanding the differences between different strategic signalling models requires good definitions of 'costs' and 'benefits'. Explicitly stating the extensive form of an interaction allows easy definition of costs and benefits. We shall see the value of these definitions in the next section. Strategic signalling has also been called conventional signalling (Maynard Smith & Harper 1988; Guilford & Dawkins 1995). We prefer to reserve the term 'conventional' for a subset of strategic signals in which the signal is tied to its meaning by a rule of convention.

Handicap versus Conventional Signals

Within the class of strategic signals, several mechanisms can maintain the evolutionary stability of variation in signal use. These differ in the source of the costs that deter universal use of the signal most likely to produce the response most desired by the signaller. That is, these signals differ in the source of the cost that prevents 'bluffing' or 'cheating'. In the case of classic handicaps, the cost is inherent in the production or use of the signal, whereas conventional signals have a cost that is determined entirely by the opponent's response to the signal. An intermediary form, the interaction handicap, has a cost imposed by the use of the signal, but this is a function of the receiver's response to the signal.

Presenting a game in the extensive form allows for a clear definition of several alternative types of signal costs. The handicap cost of signal *i*, C_i, is the difference in signaller payoffs when everything except the signal chosen is held constant. Costs may be defined relative to the cheapest signal, which we assume here to be s_1, as

$$C_i(x) = w_s(x, y, s_1, r) - w_s(x, y, s_i, r), \tag{11}$$

where x is the current signaller state, y is the current receiver state, s_i is the signal used, r is the receiver's response, and w_S is the payoff to the signaller at the end of the game as a function of player moves and states. As an example, consider the simple action–response game in Fig. 5. Here, receiver state, y, is irrelevant to the game and is thus omitted. The signaller finds itself in either state x_1 or x_2. This restricts our attention to either the left or right main branch of the tree. The signaller can use either signal s_1 or s_2 and the receiver can respond with either r_1 or r_2. If the two signals, s_1 and s_2, were equally costly to the signaller, then the payoff to the signaller when the receiver responds with r_1 would be the same regardless of which signal is given (e.g. the payoffs to the signaller at terminal nodes A and C would be equal). The same is true if the receiver responded with r_2 (resulting in equal signaller payoffs at nodes B and D). Therefore, any differences in the payoffs to the signaller at terminal nodes A and C must result from differences in inherent costs between the two signals. Note that the relative costs of the signals can differ depending on whether the signaller was in state x_1 or x_2 (the left or right hand branch of the tree). Cases in which the signal costs differ depending on which response the receiver elected to adopt are not simple handicaps in our taxonomy (q.v. interaction handicaps). Thus, the use of an extensive-form tree allows us to extract signal production costs from other cost/benefit consequences in a detailed and conditional manner, a distinct improvement over previous methods of defining such costs.

We can also calculate the consequences of a response to a signaller by comparing payoffs for different outcomes. In the action–response game shown in Fig. 5, the benefit of eliciting the response r_1 rather than r_2 is the difference between signaller payoffs at nodes A and B, or C and D, or E and F, or G and H. The signaller communicates to change the receiver's response, so we may think of this consequence as the benefit. The benefit to the signaller, of eliciting response *i*, B_i, is the difference in signaller payoffs when everything except the response is held constant. Signal benefits may be defined relative to the least preferred signal, which we assume here to be r_1, as

$$B_i(x) = w_s(x, y, s, r_i) - w_s(x, y, s, r_1). \tag{12}$$

The benefit of a signal is socially mediated; it acts through the receiver's response to the signal.

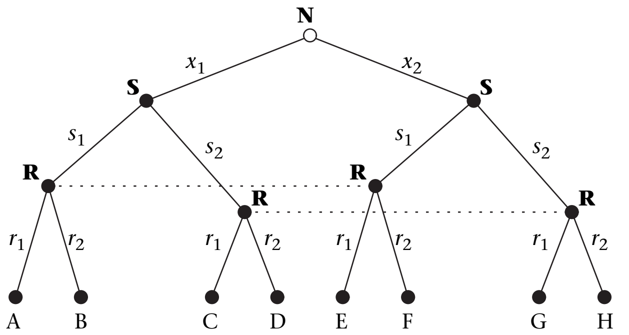


Figure 5. An extensive-form game tree of a basic action–response game. The signaller (**S**) is in one of two states, x_1 or x_2, and chooses one of two signals s_1 or s_2. The receiver (**R**) witnesses the signal and chooses one of two responses r_1 or r_2. For each of the outcomes, A–H, a payoff is specified for each of the players. For instance, at outcome A the signaller payoff will be w_S (x_1, s_1, r_1), and at outcome F the receiver will receive a payoff w_R (x_2, s_1, r_2). There is no receiver state y in the basic action–response game.

Handicapping signals

This handicap equation (11) is the inherent cost of producing or using the signal. It may include costs because of the attraction of predators or other costs because of the response of external agents whose behaviour is implicitly assumed by the model. These costs typically are a function of signaller state.

If these costs are required to maintain communication, then the signal is said to be a handicap. If communication can be maintained without these costs, then the signal is said to be conventional. The two options represent extreme cases. A distinct intermediate case, the interaction handicap, is when there is a cost to a signal, but the cost is a function of the receiver's response:

$$C_i(x, r) = w_s(x, y, s_1, r) - w_s(x, y, s_i, r) \tag{13}$$

In several games, handicaps are known to be necessary for communication to be evolutionarily stable. The vast majority of these games have one of two extensive-form game structures, either action–response games or dual-state action–response games (discussed below). The paradigm case of a handicap is the peacock's tail, undoubtedly useless, almost certainly hindering, yet preferred by females (Petrie et al. 1991) and transmitting information about mate quality (Petrie 1994). The conspicuous waste of the handicap makes the signal larger than it needs to be merely to be received.

Conventional signals

Conventional signals are those whose 'meaning' is not tied to their form. In an evolutionarily stable conventional signalling system, we can, at least theoretically, exchange two conventional signals one for the other. As long as all signallers and receivers reverse their 'meaning' simultaneously, the signalling system continues to be evolutionarily stable. An example of this is the choice of alternative song types in songbirds. In many of these systems, it is whether the song type used matches the one the opponent just used that conveys a more threatening message (Molles & Vehrencamp 2001; Vehrencamp 2001).

Conventional signals may have some small cost of production. It is important to note that it is the difference in cost between two alternative signals that determines whether they are handicapping. If the difference in cost is so small that it has no effect on the relative costs and benefits of two alternative signals, then a signal cannot be considered handicapping; for this reason Maynard Smith & Harper (2004) also use the term ‘minimal-cost signal’ to refer to what we call conventional signals. A conventional signal works upon the common interest in a subset of outcomes between signallers, even when the two players have larger conflicting interests. Because of this, conventional signals ought not to be any more conspicuous or exaggerated than necessary to convey the signal ‘message’.

The differing benefits of elicited responses maintain the stability of conventional signalling. The source of the cost that deters ‘cheating’ in a conventional signalling game is that the signal provokes a response that has negative, $B_i(x)$ (equation 12), consequences. Although the net effect is to punish bluffing signallers, this is not spiteful behaviour. The receiver is responding to the signal with the response that would lead to the highest receiver payoff if that signal had been used by a signaller playing the ESS signalling strategy.

Games in which both players signal once, then both players choose a response to the other‘s signal, may produce either conventional or handicapped solutions depending largely upon the structure of the game (see next section).

Classification of game structure

We can classify game-theoretical models of communication based upon the timing of the choices taken by the two players during the game (Table 2). The simplest game in which communication can take place requires a signaller with a variable signal, a receiver with a variable reply and that the signal be about something (i.e. a state) that is hidden from yet relevant to the receiver. In the spider mating game (Fig. 1), this state is the female’s motivational state (receptive, hungry or disinterested); in Grafen’s (1990) handicap signalling model it is the signaller’s quality, and in Maynard Smith’s (1991) Sir Philip Sidney need-signalling game, the state is the signaller’s thirst level. Action–response games are the minimal example in which these conditions are met (Hurd 1995).

Action–Response Games

We define action–response games as those games in which an actor chooses some signal, which is observed by a receiver who chooses some response based on that signal. Each of two players chooses one move: the actor chooses an action and the receiver a response. The simplest possible version of such a game, the basic action–response game, is shown in Fig. 5. The essential features are that only the signaller has a variable state, the signaller’s only move is to choose a signal as a function of that state and the receiver’s only move is to choose a response based on that signal. Most communication models in biology have this structure.

Within economics, action–response games have been named ‘signalling games’ (Gibbons 1992; Osborne & Rubenstein 1994), a name that we find too likely to lead to confusion, because this is a small subset of games in which players signal to each other. As we shall demonstrate, the results of action–response games do not generalize to other types of signalling games.

Table 2. Game structure and signal type in equilibrium strategic signalling games*

	Game structure				
Phase	Single-state action–response	Dual-state action–response	Dual-state dual-signal	Stateless mutual signalling	Mutual signalling
(Start)					
States assigned	Signaller assigned hidden state, *x*	Signaller assigned hidden state, *x* Receiver assigned hidden state, *y*	Player 1 assigned hidden state, *x* Player 2 assigned hidden state, *y*	No states assigned	Player 1 assigned hidden state, *x* Player 2 assigned hidden state, *y*
Signalling phase	Signaller chooses signal, *s*	Signaller chooses signal, *s*	Player 1 chooses signal, s' Player 2 chooses signal, s''	Player 1 chooses signal, s' Player 2 chooses signal, s''	Player 1 chooses signal, s' Player 2 chooses signal, s''
Response phase	Receiver chooses response, *r*	Receiver chooses response, *r*	Receiver chooses response, *r*	Player 1 chooses response, r' Player 2 chooses response, r''	Player 1 chooses response, r' Player 2 chooses response, r''
Payoffs	$F_{S=w(x,s,r)}$ $F_{R=w(x,r)}$	$F_{S=w(x,y,s,r)}$ $F_{R=w(x,y,r)}$	$F_{S=w(x,y,s',s'',r)}$ $F_{R=w(x,y,s'',s',r)}$	$F_{S=w(s',s'',r',r'')}$ $F_{R=w(s'',s',r'',r')}$	$F_{S=w(x,y,s',s'',r',r'')}$ $F_{R=w(x,y,s'',s',r'',r')}$
Signals at equilibrium	Handicap	Handicap	Handicap	Handicap	Conventional

*Shown are the temporal sequences of play in the five games with communication equilibrium solutions between players with conflicting interests, and the type of signal used at the equilibrium. In games in which both players choose signals, or both players choose responses, the choices are made simultaneously by the two players.

Action–response games have been used heavily within the biological literature. Although action–response games all share the same temporal sequence of events, models vary in how different states and moves are modelled. States, signals and responses may be continuous variables, or they may have some finite number of discrete alternative options. The basic discrete action–response game uses a binary state, a binary signal and a binary response (e.g. Maynard Smith 1991; Hurd 1995; Yachi 1995; Bergstrom & Lachmann 1997; Számadó 1999; Silk et al. 2000). Johnstone & Grafen (1993) model two binary states, thirst and relatedness, for a total of four signaller states, which is then signalled using a binary signal. Discrete action–response games may include more than two replies; Viljugrein (1997) modelled a game with a binary discrete state and signal, and three responses (but this reduces to a basic discrete action–response game once a strategy that is never preferred is eliminated). Continuous action–response games use the same structure as discrete action–response games, but use one or more continuously varying states or moves. For example, Grafen (1990), Godfray (1991), Johnstone & Grafen (1992b), Pagel (1993) and Reeve (1997) all treat state, signal and response as continuous variables.

Action–response games have been used to model interactions with more complicated temporal sequences. Several models of the badge of status paradigm (e.g. Owens & Hartley 1991; Johnstone & Norris 1993) analyse games with more complicated temporal structures by reducing the modelled interaction to an action–response structure. They do this by assuming that moves made later in the interaction, but not modelled explicitly, lead to some specified final payoffs for the players.

Dual-state Action–Response Games

The second most widely used signalling game structure is a slightly more complicated version of the action–response game. These games add a receiver state to an action–response game structure. Virtually all of these models have been 'Sir Philip Sidney games', investigating the effect of relatedness on signals of need (e.g. Johnstone & Grafen 1992a; Maynard Smith 1994; Bergstrom & Lachmann 1998; Lachmann & Bergstrom 1998), although one model (Adams & Mesterton-Gibbons 1995) investigates threat-display use when one signaller threatens and the other responds. All these dual-state action–response games also require handicapping costs to stabilize communication between players with conflicting interests.

These games no longer have action–response game structures. Their structures do however collapse in practice across receiver states to a single action–response game against a single expected receiver type. This is because the signaller never gains any information about the receiver state, so there is never any opportunity for the signaller to do anything but treat the receiver population as a single weighted average playing a mixed strategy (even though the response is in fact based on the unknowable receiver state) (Maynard Smith 1994). This game may produce 'dishonest' signalling (Johnstone & Grafen 1992a) and cost-free signalling (Bergstrom & Lachmann 1998) through 'pooling', discussed further in the section Action–Response Games and Handicaps.

Dual-state, Dual-signal Games

Maynard Smith (1994) presents a version of the Sir Philip Sidney game (Fig. 6) in which both players have hidden states, and both players signal to each other, but only one player moves after the signals. The signals in this game are not simultaneous.

Mutual Signalling Games

The earliest biological example of an ESS signalling game (Enquist 1985) is based on a relatively complicated structure. In this game, each player has a hidden state, and each player chooses a signal, then observes the opponent's signal before choosing a move in an end game. We have referred to this structure as the conventional signalling game, because it is the simplest game that produces a conventional signalling equilibrium between players with conflicting interests (Hurd & Enquist 1998). The successive moves by the same player allows for the mitigation of negative consequences and capitalization on positive consequences using later moves in the same game; this makes conventional signalling possible (Hurd & Enquist 1998).

We define mutual signalling games as those in which both players signal, and both players respond to signals. The players are referred to as 'player 1' and 'player 2', or 'ego' and 'opponent', rather than 'signaller' and 'receiver'. This form of interaction, in which both players signal, and both players react to their opponent's signal, more closely resembles an agonistic interaction between animals. Models with this structure are becoming more widely used (e.g. Számadó 2000, 2003). The weakness of these games is that the players are assumed to choose their signals simultaneously.

Stateless Mutual Signalling Game

One influential signalling game in the biological literature models communication between players without hidden states (Kim 1995). This stateless signalling game is a Hawk–Dove game preceded by a signalling phase. Both players begin by choosing one of two signals, m_0 or m_1 (see Fig. 7), then use patterns of signal matching and mismatching to choose moves in a subsequent Hawk–Dove end game (Maynard Smith & Parker 1976; Maynard Smith 1982).

This game is unique among those reviewed here in that the players do not have variable states, or at least they do not have payoff-relevant states. Instead, the signals provide information indicating the relative probability of alternative future acts. The signalling equilibrium in games with this structure is fragile to violations of several key assumptions, discussed below.

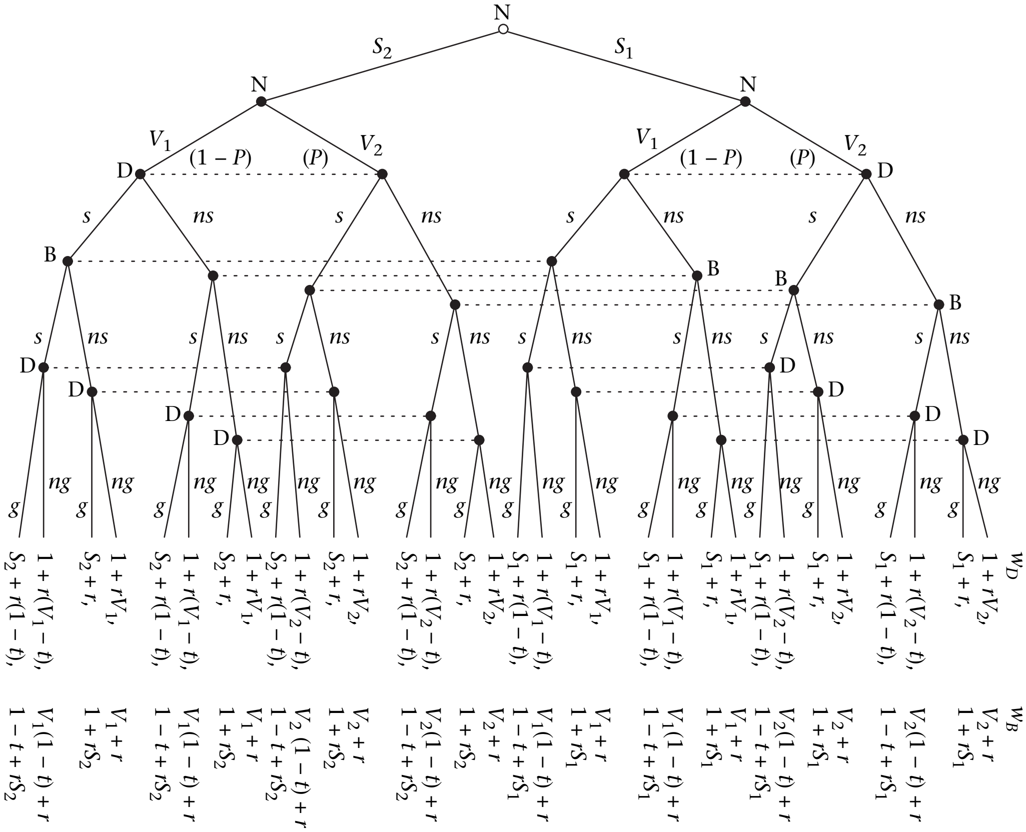


Figure 6. Maynard Smith's (1994) mutual signalling version of the Sir Philip Sidney game. All variable names are as in the original version, donor states are S_1 and S_2, beneficiary states are V_1 and V_2, beneficiary moves are signal (*s*) or no signal (ns), donor moves are give (g) or do not give (ng). To see how the extensive form differs from an intuitive description of a game, compare this figure with Fig. 2 in Maynard Smith (1994).

Game structure and signal type

The structure of a game influences the types of signal that can exist at a communication equilibrium. Table 2 summarizes the temporal structures of the games reviewed here, and the types of signals that can exist at communication equilibria. Although the simplest games require handicaps for signal use to be evolutionarily stable, more complicated games do not. The games that produce conventional signals are by no means overly complex. They remain a caricature of real interactions, but they are much closer to biological reality than are action–response games in that both players signal and react to signals. Adding a handicap to a stable conventional signalling game produces an effect opposite to the predictions of the handicap principle, in that the costly signal is used by the weaker signaller, not the stronger one (Hurd 1997b). This effect means that we cannot study only the simplest games and extrapolate their results to cover interactions with more involved temporal structure.

Action–Response Games and Handicaps

Action–response games may produce conventional signalling equilibria when the players have common interests, but they produce handicapped signalling equilibria only when the players have conflicting interests (Spence 1973; Grafen 1990; Godfray 1991; Maynard Smith 1991; Hurd 1995). Interests are common when each player's ranking of payoffs is the same. Borderline cases, in which common and conflicting interests are difficult to classify, arise when signallers 'pool' signals (i.e. when signallers in different states use the same signal; Gibbons 1992).

For example, in the spider game (Fig. 1), the male receiver has common interests with disinterested and receptive females and conflicting interests with hungry females. Hungry and receptive females pool to a single signal, however, so the male is forced to choose a strategy that responds one way to disinterested females and another way to both hungry and receptive females. The male must treat hungry and receptive females as one average class of 'females who use signal C', because he has no way of ever separating responses between the two component classes. If hungry and receptive females are averaged in this way, then the males have common interests with both the G-using and averaged C-using females, and a costless signal may be used at a signalling equilibrium. The signal is simply ambiguous. It seems inappropriate to call such a signalling system 'deceptive' or 'dishonest', or even 'honest on average'. These pseudoconflicting interests with pooled signaller types underly cost-free signalling in action–response games (e.g. Silk et al. 2000).

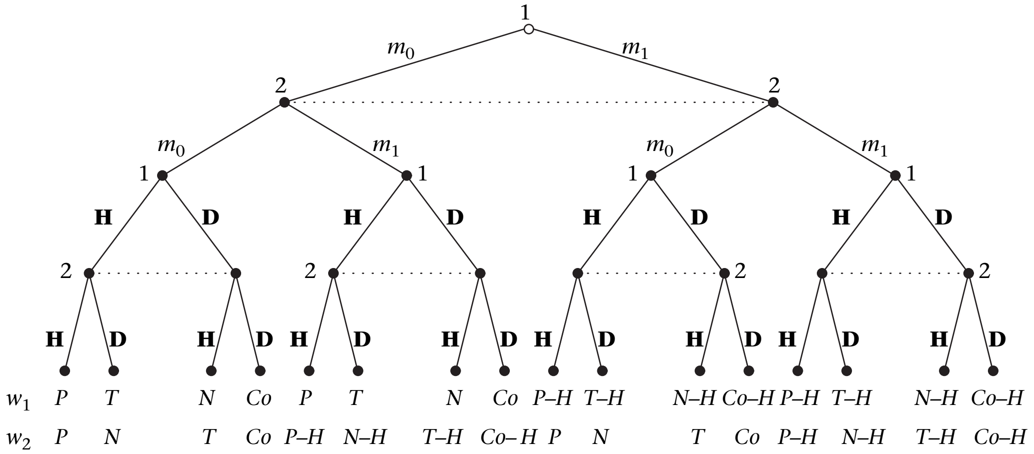


Figure 7. Kim's (1995) stateless signalling game. Payoffs for the Hawk–Dove (H–D) subgames have been replaced by the generic variables T, Co, N, P (temptation, coordination, neutral and punishment); the results are identical to the traditional Hawk–Dove payoffs (with V and C). A signalling ESS exists when $T > \text{Co} \geq N > P$. Generic payoffs for Hawk–Dove players are equivalent to: $T = V$, $\text{Co} = V/C$, $N = 0$, $P = (V - C)/2$ (where the generic payoffs are on the left, and classical Hawk–Dove payoffs are on the right). H is the handicap cost of using signal m_1. The signals in this game are effectively simultaneous. The second player chooses the signal to be used before learning of the first player's choice of signal.

Another borderline example is the Johnstone & Grafen (1993) version of the Sir Philip Sidney game (Fig. 8). This game models two different types of signaller (differing in relatedness to the receiver), which are then assigned to two levels of thirst. This arrangement decomposes to a four-state (type × thirst level) action–response game. The receiver receives no information other than the signal, so each of the types then collapses into an average expected type according to the signal chosen (Maynard Smith 1994). This game then simplifies into a single-state action–response game. In other words, these two models, the Sir Philip Sidney game and the action–response game, are really the same thing.

Mutual Signalling Games and Signal Type

The averaging out of opponent types through signal pooling will occur whenever players cannot modify behavioural choices with later moves. The single-state and dual-state action–response games each allow only one move per player. The stateless signalling and conventional signalling games comprise a signalling phase and a response phase. The response phase allows players to modify their initial choice, the signal choice, given subsequently gained information. Players may choose behaviours that will either maximize benefit in best cases, or minimize costs in worse cases. This fine-tuning provides a mechanism for types that were pooled earlier in the game to be subsequently separated. This level of complexity allows for conventional signalling between players with conflicting interests (Hurd & Enquist 1998).

Stateless Mutual Signalling Games and Handicaps

Stateless signalling games (Fig. 7) begin with an initial mixed ESS signalling move. With no underlying state, the signalling move must be a choice made between

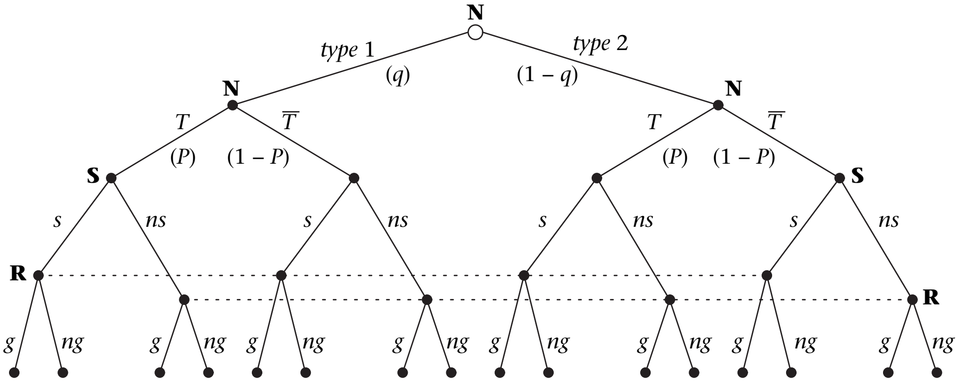


Figure 8. The 'Sir Philip Sidney' game as modelled by Johnstone & Grafen (1993). A signaller (**S**, the beneficiary) is of one of two types, either closely related (type 1), or less closely related (type 2), chosen in a move by nature (**N**). These signallers are either thirsty (T) or not ($\overline{T}$), and signal (s) or do not (ns). The receiver (**R**, the donor) then either gives (g) the resource to the beneficiary, or does not (ng). The receiver has no way of ever learning whether the signaller is of type 1 or type 2, so the distinction between them can be ignored and all signallers treated as one weighted average type. This reduces the game to a basic action–response game.

alternatives with equal expected net benefits. If interests conflict then some moves that follow the signal must lead to an outcome more beneficial to one player than the other. Any signal used with the expectation of increasing the probability of a preferable outcome must have a balancing cost to negate the benefit. In the case of Kim's (1995) aggressiveness signalling game (and any other game without state-dependent signalling, and having an end game with the broad properties of the Hawk–Dove game) the signal that benefits the signaller most must cost more through some wasteful balancing cost (P. L. Hurd, unpublished data). If one signal were to lead to a higher net payoff, all signallers would choose that one signal. A handicap is required to negate the benefits of the 'better' signal so that all signals payoff equally in the end.

Discussion

Explaining the breadth and forms of communication between animals is a classic topic within evolutionary biology. It is a subject that has, however, suffered from ambiguous definition, and intuitive, vague, verbal models. Significant advances in understanding have resulted largely from lucid formal treatments (e.g. Maynard Smith & Harper 1988; Grafen 1990; Hasson 1997). In this paper, we have offered a strategic taxonomy of communication systems that naturally classifies signalling according to both (1) signal properties and (2) known dynamics theoretically capable of maintaining signalling at an evolutionary equilibrium (regardless of whether the system is at equilibrium, the properties of the dynamics can still serve in classification). This strategic analysis of signals and communication systems is based upon the extensive-form game representation of interactions, which matches the time flow of interactions between individuals.

For systems in a signalling equilibrium, we distinguished between those cases in which all players have access to all signals (strategic signalling) and cases in which they do not (performance signalling). Performance signalling is a sufficient mechanism for maintaining evolutionarily stable communication (e.g. the sequential assessment game: Enquist & Leimar 1983; Leimar & Enquist 1984). In the case of strategic signalling, alternative signals must differ in how costly they are to use if communication is to be evolutionarily stable. In turn, this requires that different signals be optimal for different types of senders. We have limited the use of the term 'handicap' to those cases when the cost is associated with the signal, and not the response to the signal (as is the case with conventional signalling). Thus, with this interpretation, the handicap is something inherent to the form of the signal.

Vehrencamp (2000) presented the classification system that most closely resembles the one proposed here. Both systems focus on the source of signal costs as the critical character. The chief differences between this signal taxonomy and Vehrencamp's are that (1) we have formalized the definition of signal cost, and (2) we focus more on game structure rather than biological interpretation. The two systems divide conventional from handicapped models in the same way. We prefer the term 'interaction handicap' to Vehrencamp's 'vulnerability handicap', because the concept of vulnerability seems tuned specifically to the example of threat displays, and it distracts from the key property of an interaction between signal property and receiver response. Our classic handicap category includes both Vehrencamp's 'quality handicap' and 'general handicap'. The differences between these two categories in terms of our framework is again clearest in the case of aggressive interactions. Quality handicaps are those whose cost component is a function of 'fighting ability, stamina, health, condition' (Vehrencamp 2000, page 279), factors for which signallers in differing states pay different costs for a fight. General handicaps are those whose costs are a function of 'motivation, need, resource valuation' (Vehrencamp 2000, page 279), factors for which differing state signallers pay equal costs for a fight, but expect different benefits. Although the distinction between quality and general handicaps could be formalized in the present system, there are two reasons for not doing so: (1) there is no clear distinction between quality and need states in generic games, and more importantly, (2) we see no reason to expect that the two types of handicap would show different properties.

Maynard Smith & Harper (2004) have developed a classification of signalling similar to the one presented in this paper. Their system draws on semiotics and refers to symbols, icons and indexes. In semiotics, a symbol is a sign that does not resemble the thing that it represents or expresses. Words and numerals are typical human symbols. The relationship between the form and meaning of a symbol is arbitrary. Signals used solely for species identification would be a biological example of a symbol. An icon is a sign that represents its object by its similarity to it. The relation between form and meaning of an icon is not arbitrary. An example could be a threat display that includes some intention movement elements of attack, like pointing the beak towards another individual (e.g. Stokes 1962; Dunham 1966; Blurton Jones 1968). Maynard Smith & Harper's icon and symbol categories are difficult to translate into our taxonomy. Issues such as similarity of form (or lack thereof) have no meaning to game theory or to equilibrium solutions in general. Semiotic terminology focuses attention on the signal form rather than on strategic issues (i.e. what makes a signalling system evolutionarily stable). For example, semiotic treatments cannot address the biologically critical issue of signal cost.

Several other strategic classification systems for biological signals exist, but are not particularly well suited to drawing attention to the distinctions that are central to this work. Hasson (1994) categorized signalling systems by whether they enable the receiver to make a discrimination and whether the signaller and receiver are attempting to accomplish a mutual goal (which, as Maynard Smith & Harper (1995) pointed out, are not mutually exclusive categories). These categorizations all depend upon comparing the payoffs for various outcomes without examining the effect of the signal per se on payoff. Maynard Smith & Harper (1995) also included costs of receiver responses to signals in their accounting of signal cost. Although their taxonomy also attempts to classify signalling systems according to the stabilizing mechanism, it does not separate the effects of signals on receivers' actions from the other ramifications of the signal. This distinction is critical when examining conventional signalling.

Another approach is to classify according to the manner in which the receiver interprets the signal. Payne & Pagel (1997) presented a classification based upon the receiver criterion in responding to a set of repeated graded displays. For instance, consider a signalling system in which the receiver judges a best response as a function of the mean value of the signals used thus far by a signaller. This system is fundamentally different from the one in which the receiver judges how to respond as a function of the signal with the greatest magnitude used thus far. The receiver's response criterion, however, must be determined by the stabilizing mechanism. Differing mechanisms may therefore exist in which the receiver's decision criterion is identical, so classification by the stabilizing mechanism will be more precise. Payne & Pagel's (1997) system also works only for signalling systems in which a single type of graded signal is used repeatedly. The taxonomy has nothing to contribute when the display repertoire consists of several discretely different displays.

Finally, communication is sometimes classified as either honest or dishonest (reviews in Johnstone 1997; Bradbury & Vehrencamp 1998). All signalling systems out of equilibrium may be seen as dishonest, because the receiver does not correctly anticipate the signaller's actions. When in evolutionary equilibrium, on the other hand, 'dishonest' communication merely means that a signaller has chosen a move from within one of the receiver's information sets that the receiver was not betting on. To call this sort of signalling dishonest is akin to calling losing lottery tickets 'dishonest'. The honest/dishonest dichotomy glosses over most signalling systems that appear to be ambiguous but informative (Hurd & Enquist 2001). In these systems, the signals provide some information about the signaller's impending acts without specifying what they are certain to be.

In conclusion, extensive-form game theory provides an explicit language for defining types of biological signals and classifying the temporal structure of signalling interactions. These organizing principles have proved contentious issues in the history of the study of biological communication, in no small part because of a lack of common definitions for intuitive terms. Game theory has the ability to express the ideas central to hermeneutic theory such as meaning and message, as well as critical concepts that hermeneutics is blind to, such as handicap and conventional signals. Table 1 provides an overview of the classes within our proposed taxonomy, along with a few examples. The 'out of equilibrium' category is the least developed, partly because theories of evolutionary dynamics at disequilibria are relatively underdeveloped (with the notable exception of sexually selected signals). Several results indicate, however, that this may be a common situation, in particular when signalling options are not limited to a few signals or dimensions (see e.g. Enquist et al. 2002). Our taxonomy of signalling focuses on strategies of social interaction while ignoring the problem of signalling transmission as well as the topic of signaller constraints. It is obvious that signals used in long-range communication must be exaggerated to reach the receiver independent of whether a conflict is present or whether a handicap is required for evolutionary stability. Natural history factors will always remain unabstractable in the face of biological theory.

Acknowledgments

We thank John Maynard Smith, David Harper, Sandy Vehrencamp and Eldridge Adams for valuable discussions about these issues and topics, as well as the helpful advice and constructive criticism of George Uetz and two anonymous referees (one referee suggested Gintis (2003) as a practical introduction to extensive-form models; we have previously suggested Gibbons (1992); each has its shortcomings). We thank Chris Sturdy, Eric Verbeek, Walter Espinoza and Heather Hinam for their comments on earlier versions of this manuscript. This work was funded by a grant from the Swedish Research Council to M.E. and a Natural Sciences and Engineering Research Council (Canada) Discovery grant to P.L.H.

Appendix: Definitions

The definitions for some of these terms may not generalize well outside the context of this paper.

Communication Terms

We define communication in two partly complementary ways, operational and state definitions. The operational definition focuses on behaviour and is the easiest to use empirically. The state definition focuses on internal changes in the receiver. A third definition, information transfer, is similar to the state definition, but it focuses explicitly on information.

Operational definition

Communication is a change in receiver behaviour in response to a signal. If different signals cause different responses, then communication has taken place. This is how communication is identified empirically.

State definition

Communication is a change of the receiver's internal state in response to a signal. Internal state refers to memory or other states within the nervous system. Such changes in internal state may be expressed in behaviour, either immediately or later, but this is not necessary. For example, in the case of sexual imprinting, a casual observer of the signal may not perceive the behavioural response required to meet the operational definition of communication. However, the receiver clearly has been changed internally, as evidenced operationally years later when sexual behaviour is displayed. According to this definition, communication may change what we think, even if it is difficult to show an obvious change in behaviour.

Information transfer definition

Communication is a transfer of information. If the information is reliable (most obviously when signalling behaviour is evolutionarily stable), then this interpretation makes sense. In this situation, the sender makes relevant information available to the receiver by using different signals in different states.

Message

The implication that a signal makes about signaller states. If communication can be interpreted as a transfer of information, then the message is the information that a signal conveys (Smith 1977, page 18).

Meaning

The interpreted significance of the message to the particular receiver, which involves the appropriate response (Smith 1977, page 19). The meaning of a signal is what the signal 'means' that the receiver ought to do in response. The meaning of a signal at equilibrium may be simply anthropomorphized (e.g. message: 'aposematic display: probably toxic', meaning: 'Leave alone!'). When communication is out of equilibrium, examples may be impossible to interpret in this way.

Honesty

We see several different possible definitions for the terms 'honesty' and 'dishonesty'. (1) Honesty is when the sender provides all relevant information to the receiver. Doing so allows the receiver to infer the signaller's state from the signal chosen without ambiguity. For instance, in the spider game, the female does not inform the male whether she is receptive or hungry, only that she is one of the two. Johnstone & Grafen (1993) would call the female spider 'honest on average', because she provides ambiguous information, but information that benefits the male if he responds to it. We think that it is inappropriate to classify the spider game as either 'honest', 'dishonest' or 'honest on average'. The signaller provides information such that the receiver has an accurate, but probabilistic, knowledge of signaller states after receiving the signal. This concept of honesty applies only to communication at an evolutionary equilibrium. (2) A related concept, dishonesty, occurs when a signaller uses a signal in a way not anticipated by the receiver. For example, an imperfect mimic may avoid being attacked by predator by resembling an unprofitable prey. Such dishonesty can exist only if strategies are out of equilibrium (i.e. the signaller has not evolved the ability to discriminate the imperfect mimic from the model). We dislike the term 'honest' and suggest focusing instead on the extent to which signalling is informative.

Game Theory Terms

Extensive form

An explicit representation of a game as a series of decision nodes and vertices (Fudenberg & Tirole 1991; Gibbons 1992) that allows explicit depiction of a game with a temporal sequence and ambiguous information. Extensive-form models allow behaviours to be chosen contingent upon an individual's underlying states, and/or events that occurred earlier in the game.

Information set

A set of nodes in an extensive-form game between which a player cannot distinguish. One of the nodes within the information set represents the true state of play of the game. A player must have a 'belief' about the probability that each node within the set is the true state of play. In the spider game, two information sets, one for each signal, join the female states (see also Communication as transfer of information).

Hidden state

A state that presumably has a relevant payoff to a player, but which is unknown to the player whose move it is. States that differ between nodes contained within an information set are hidden states.

Strategy

A set of rules that describes what move should be made under all possible situations. A strategy for an extensive-form game prescribes a move for each node or information set that a player may encounter.

Beliefs

The distribution of probability assigned to each of the nodes within an information set for a player. Beliefs may be described as correct or incorrect. If the player's beliefs match the true stochastic probabilities that the nodes are reached during play, then the belief is correct.

Preference ranking and conflict

The order in which a player ranks different outcomes. If, for example, an extensive-form game has four end points (a, b, c, d), and one player has the preference ranking of $a > b > c > d$ and another has the preference ranking $c > b > a > d$, then the payers have conflicting interests.

Types of Signals

Here we briefly review the types of signals defined for evolutionarily stable signalling systems.

Cost-free or costless signal

A signal whose cost of production is zero.

Conventional signals

A set of signals whose meanings are independent of their form. A and B are conventional signals if both the sender's choice of signal and the receiver's responses can be exchanged without significantly changing payoff. These contingencies require that all signals within the set are cost-free or equally costly to produce.

Handicap

The increase in cost associated with using one signal rather than another.

Production handicap

An inherent cost of production or use of a signal. This type of handicap is paid regardless of the receiver's reaction to the signal.

Interaction handicap

Signalling cost of responses from the receiver that differ between signals because of the nature of the differences between signals. The magnitude of such handicaps may vary between individuals. For instance, one signal may have a form that exposes the signaller more to counterattack than does an alternative signal. Some individuals, those of high resource-holding potential, may face smaller costs for this handicap, but there is no cost to any player unless the opponent decides to attack immediately following the display.

Performance displays

A signal whose use is limited to a subset of signallers. This sort of signal is reliable because it is impossible for some senders to produce the display. Two types of performance displays are index displays and information-constrained displays.

Index displays

A performance signal whose use is constrained by physiology or physical ability.

Information-constrained displays/signals

A performance signal whose use is limited to a subset of signallers because of informational constraints.

References

Adams, E. S. & Mesterton-Gibbons, M. 1995. The cost of threat displays and the stability of deceptive communication. *Journal of Theoretical Biology*, **175**, 405–421.

Arak, A. & Enquist, M. 1995. Conflict, receiver bias and the evolution of signal form. *Philosophical Transactions of the Royal Society of London, Series B*, **349**, 337–344.

Baerends, G. P. & Baerends-van Roon, J. 1950. An introduction to the study of the ethology of the cichlid fishes. *Behaviour, Supplement*, **1**, 1–242.

Basolo, A. L. 1990. Female preference predates the evolution of the sword in swordtail fish. *Science*, **250**, 808–810.

Bergstrom, C. T. & Lachmann, M. 1997. Signalling among relatives. I. Is costly signalling too costly? *Philosophical Transactions of the Royal Society of London, Series B*, **352**, 609–617.

Bergstrom, C. T. & Lachmann, M. 1998. Signalling among relatives. III. Talk is cheap. *Proceedings of the National Academy of Science, U.S.A.*, **95**, 5100–5105.

Blurton Jones, N. G. 1968. Observations and experiments on causation of threat displays of the great tit (*Parus major*). *Animal Behaviour Monograph*, **1**, 75–158.

Bradbury, J. W. & Vehrencamp, S. L. 1998. *Principles of Animal Communication*. Sunderland, Massachusetts: Sinauer.

Caryl, P. G. 1979. Communication by agonistic displays: what can games theory contribute to ethology? *Behaviour*, **68**, 136–169.

Chapman, T., Arnqvist, G., Bangham, J. & Rowe, L. 2003. Sexual conflict. *Trends in Ecology and Evolution*, **18**, 41–47.

Dawkins, R. & Krebs, J. R. 1978. Animal signals: information or manipulation? In: *Behavioural Ecology: an Evolutionary Approach*. 1st edn (Ed. by J. R. Krebs & N. B. Davies), pp. 282–309. Oxford: Blackwell.

Dittrich, W., Gilbert, F., Green, P., McGregor, P. & Grewcock, D. 1993. Imperfect mimicry: a pigeons perspective. *Proceedings of the Royal Society of London, Series B*, **251**, 195–200.

Dunham, D. W. 1966. Agonistic behavior in captive rose-breasted grosbeaks, *Pheucticus ludovicianus* (L.). *Behaviour*, **27**, 160–173.

Enquist, M. 1985. Communication during aggressive interactions with particular reference to variation in choice of behaviour. *Animal Behaviour*, **33**, 1152–1161.

Enquist, M. & Jakobsson, S. 1986. Decision making and assessment in the fighting behaviour of *Nannacara anomala* (Cichlidae, Pisces). *Ethology*, **72**, 143–153.

Enquist, M. & Leimar, O. 1983. Evolution of fighting behaviour: decision rules and assessment of relative strength. *Journal of Theoretical Biology*, **102**, 387–410.

Enquist, M., Leimar, O., Ljungberg, T., Mallner, Y. & Segerdahl, N. 1990. A test of the sequential assessment game: fighting in the cichlid fish *Nannacara anomala*. *Animal Behaviour*, **40**, 1–14.

Enquist, M., Arak, A., Ghirlanda, S. & Wachtmeister, C.-A. 2002. Spectacular phenomena and limits to rationality in genetic and cultural evolution. *Philosophical Transactions of the Royal Society of London, Series B*, **357**, 1585–1594.

Eshel, I. 1983. Evolutionary and continuous stability. *Journal of Theoretical Biology*, **103**, 99–111.

Eshel, I. 1996. On the changing concept of evolutionary population stability as a reflection of a changing point of view in the quantitative theory of evolution. *Journal of Mathematical Biology*, **18**, 123–133.

Fisher, R. A. 1930. *The Genetical Theory of Natural Selection*. 1st edn. Oxford: Clarendon.

Fudenberg, D. & Tirole, J. 1991. *Game Theory*. Cambridge, Massachusetts: MIT Press.

Gerhardt, H. C. 1994. The evolution of vocalization in frogs and toads. *Annual Review of Ecology and Systematics*, **25**, 293–324.

Geritz, S. A. H., Kisdi, É., Meszéna, G. & Metz, J. A. J. 1998. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology*, **12**, 35–57.

Gibbons, R. 1992. *A Primer in Game Theory*. London: Harvester Wheatsheaf.

Gintis, H. 2003. *Game Theory Evolving: a Problem Centered Introduction to Modeling Strategic Interaction*. Princeton, New Jersey: Princeton University Press.

Godfray, H. C. J. 1991. Signalling of need by offspring to their parents. *Nature*, **352**, 328–330.

Grafen, A. 1990. Biological signals as handicaps. *Journal of Theoretical Biology*, **144**, 517–546.

Guilford, T. & Dawkins, M. S. 1995. What are conventional signals? *Animal Behaviour*, **49**, 1689–1695.

Gullion, G. W. 1952. The displays and calls of the American coot. *Wilson Bulletin*, **64**, 83–97.

Hasson, O. 1994. Cheating signals. *Journal of Theoretical Biology*, **167**, 223–238.

Hasson, O. 1997. Towards a general theory of biological signaling. *Journal of Theoretical Biology*, **185**, 139–156.

Houston, A. I. & McNamara, J. M. 1999. *Models of Adaptive Behaviour, an Approach Based on State*. Cambridge: Cambridge University Press.

Hurd, P. L. 1995. Communication in discrete action–response games. *Journal of Theoretical Biology*, **174**, 217–222.

Hurd, P. L. 1997a. Game theoretical perspectives on conflict and biological communication. Ph.D. thesis, Stockholm University.

Hurd, P. L. 1997b. Is signalling of fighting ability costlier for weaker individuals? *Journal of Theoretical Biology*, **184**, 83–88.

Hurd, P. L. & Enquist, M. 1998. Conventional signalling in aggressive interactions: the importance of temporal structure. *Journal of Theoretical Biology*, **192**, 197–211.

Hurd, P. L. & Enquist, P. L. 2001. Threat display in birds. *Canadian Journal of Zoology*, **79**, 931–942.

Jakobsson, S., Radesäter, T. & Järvi, T. 1979. On the fighting behaviour of *Nannacara anomala* (Pisces, Cichlidae) males. *Zeitschrift für Tierpsychologie*, **25**, 257–314.

Johnstone, R. A. 1997. The evolution of animal signals. In: *Behavioural Ecology: an Evolutionary Approach*. 4th edn (Ed. by J. R. Krebs & N. B. Davies), pp. 155–178. Oxford: Blackwell.

Johnstone, R. A. 1998. Game theory and communication. In: *Game Theory and Animal Behavior* (Ed. by L. A. Dugatkin & H. K. Reeve), pp. 64–93. New York: Oxford University Press.

Johnstone, R. A. & Grafen, A. 1992a. The continuous Sir Philip Sidney game: a simple model of animal signalling. *Journal of Theoretical Biology*, **156**, 215–234.

Johnstone, R. A. & Grafen, A. 1992b. Error-prone signalling. *Proceedings of the Royal Society of London, Series B*, **248**, 229–233.

Johnstone, R. A. & Grafen, A. 1993. Dishonesty and the handicap principle. *Animal Behaviour*, **46**, 759–764.

Johnstone, R. A. & Norris, K. 1993. Badges of status and the cost of aggression. *Behavioral Ecology and Sociobiology*, **32**, 127–134.

Kim, Y.-G. 1995. Status signalling games in animal contests. *Journal of Theoretical Biology*, **176**, 221–231.

Krebs, J. R. & Dawkins, R. 1984. Animal signals: mind-reading and manipulation. In: *Behavioural Ecology: an Evolutionary Approach*. 2nd edn (Ed. by J. R. Krebs & N. B. Davies), pp. 380–402. Sunderland, Massachusetts: Sinauer.

Lachmann, M. & Bergstrom, C. T. 1998. Signalling among relatives. II. Beyond the tower of Babel. *Theoretical Population Biology*, **54**, 146–160.

Lande, R. 1981. Models of speciation by sexual selection on polygenic traits. *Proceedings of the National Academy of Sciences, U.S.A.*, **78**, 3721–3725.

Leimar, O. & Enquist, M. 1984. The effects of asymmetries in owner–intruder interactions. *Journal of Theoreteical Biology*, **111**, 475–491.

Martin, W. F. 1972. Evolution of vocalization in the genus *Bufo*. In: *Evolution in the Genus Bufo* (Ed. by W. F. Blair), pp. 279–309. Austin: University of Texas Press.

Maynard Smith, J. 1974. Theory of games and the evolution of animal contests. *Journal of Theoretical Biology*, **47**, 209–221.

Maynard Smith, J. 1982. *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.

Maynard Smith, J. 1991. Honest signalling: the Philip Sidney game. *Animal Behaviour*, **42**, 1034–1035.

Maynard Smith, J. 1994. Must reliable signals always be costly? *Animal Behaviour*, **47**, 1115–1120.

Maynard Smith, J. & Harper, D. G. C. 1988. The evolution of aggression: can selection generate variability? *Philosophical Transactions of the Royal Society of London, Series B*, **319**, 557–570.

Maynard Smith, J. & Harper, D. G. C. 1995. Animal signals: models and terminology. *Journal of Theoretical Biology*, **177**, 305–311.

Maynard Smith, J. & Harper, D. G. C. 2004. *Animal Signals*. Oxford: Oxford University Press.

Maynard Smith, J. & Parker, G. A. 1976. The logic of asymmetric contests. *Animal Behaviour*, **24**, 159–175.

Molles, L. E. & Vehrencamp, S. L. 2001. Songbird cheaters pay a retaliation cost: evidence for auditory conventional signals. *Proceedings of the Royal Society of London, Series B*, **268**, 2013–2019.

Munn, C. A. 1986. Birds that 'cry wolf'. *Nature*, **319**, 143–145.

Osborne, M. J. & Rubenstein, A. 1994. *A Course in Game Theory*. Cambridge, Massachusetts: MIT Press.

Owens, I. P. F. & Hartley, I. R. 1991. 'Trojan sparrows': evolutionary consequences of dishonest invasion. *American Naturalist*, **138**, 1187–1205.

Pagel, M. 1993. Honest signalling among gametes. *Nature*, **363**, 539–541.

Parker, G. 1979. Sexual selection and sexual conflict. In: *Sexual Selection and Reproductive Competition in Insects* (Ed. by M. S. Blum & N. A. Blum), pp. 123–166. New York: Academic Press.

Payne, R. J. H. & Pagel, M. 1997. Why do animals repeat displays? *Animal Behaviour*, **54**, 109–119.

Petrie, M. 1994. Improved growth and survival of offspring of peacocks with more elaborate trains. *Nature*, **371**, 598–599.

Petrie, M., Halliday, T. & Sanders, C. 1991. Peahens prefer peacocks with elaborate trains. *Animal Behaviour*, **41**, 323–331.

Reeve, H. K. 1997. Evolutionarily stable communication between kin: a general model. *Proceedings of the Royal Society of London, Series B*, **264**, 1037–1040.

Selten, R. 1980. A note on evolutionarily stable strategies in asymmetric animal contests. *Journal of Theoretical Biology*, **84**, 93–101.

Silk, J. B., Kaldor, E. & Boyd, R. 2000. Cheap talk when interests conflict. *Animal Behaviour*, **59**, 423–432.

Smith, W. J. 1977. *The Behavior of Communicating: an Ethological Approach*. Cambridge, Massachusetts: Harvard University Press.

Spence, A. M. 1973. Job market signalling. *Quarterly Journal of Economics*, **87**, 355–374.

Stokes, A. W. 1962. Agonistic behaviour among blue tits at a winter feeding station. *Behaviour*, **19**, 118–138.

Számadó, S. 1999. The validity of the handicap principle in discrete action–response games. *Journal of Theoretical Biology*, **198**, 593–602.

Számadó, S. 2000. Cheating as a mixed strategy in a simple model of aggressive communication. *Animal Behaviour*, **59**, 221–230.

Számado, S. 2003. Threat displays are not handicaps. *Journal of Theoretical Biology*, **221**, 327–348.

Trivers, R. L. 1974. Parent–offspring conflict. *American Zoologist*, **14**, 249–264.

Vehrencamp, S. L. 2000. Handicap, index, and conventional signal elements of bird song. In: *Animal Signals: Signalling and Signal Design in Animal Communication* (Ed. by E. Espmark, T. Amundsen & G. Rosenqvist), pp. 277–300. Trondheim: Tapir.

Vehrencamp, S. L. 2001. Is song-type matching a conventional signal of aggressive intentions? *Proceedings of the Royal Society of London, Series B*, **268**, 1637–1642.

Viljugrein, H. 1997. The cost of dishonesty. *Proceedings of the Royal Society of London, Series B*, **264**, 815–821.

Wiley, R. H. 1983. The evolution of communication: information and manipulation. In: *Animal Behaviour. Vol. 2: Communication* (Ed. by T. R. Halliday & P. J. B. Slater), pp. 156–189. Oxford: Blackwell.

Williams, G. C. 1966. *Adaptation and Natural Selection: a Critique of Some Current Evolutionary Thought*. Princeton, New Jersey: Princeton University Press.

Yachi, S. 1995. How can honest signalling evolve? The role of the handicap principle. *Proceedings of the Royal Society of London, Series B*, **262**, 283–288.