Honest Signaling

P. L. Hurd, University of Alberta, Edmonton, AB, Canada

Introduction

Classical ethologists largely thought of communication in terms of senders broadcasting information indicating their internal motivational state, and receivers upon whom these signals then acted as releaser stimuli. The receiver's response to these signals was assumed to be a behavioral response appropriate to the signaler's state. The outcome of this interaction would then be of mutual benefit to both signaler and receiver. Evolution was thought to act to make signals less ambiguous about the signaler state, and more efficient and effective at longer and longer ranges. Unfortunately, the logical basis of this classical approach has a potential flaw.

The outline of a challenging problem to this classical view emerged during the 1970s with the rise of explicitly 'selfish' gene-centric views of evolution. The application of game theoretical thinking to social behavior highlighted the importance of conflicting interests in signaling interactions. Most conspicuous communication occurs between individuals in some sort of conflict. Fighting animals often communicate their strength to one another. A prospective suitor will show-off to a prospective mate using courtship displays. Dependent young beg their parents for food. If honest, then all these situations seem to involve individuals giving away information to their disadvantage: some individuals informing a receiver that they are weak, unfit mates, or less hungry than their siblings. Such submaximally escalated signals ought to encourage the opponent to continue fighting, the wooed to spurn, parents to provision less. On the other hand, if all fighting animals signal that they are incredibly strong, suitors always signal that they are the most conceivably deserving mates, and offspring signal that they are starving then, presumably the receivers of these signals would evolve to ignore them entirely.

Do Displays Transfer Information?

An early hypothesis advanced to solve the dilemma of communication with conflicting interests was to question whether signals actually conveyed any information at all. The prediction that there ought to be no information transmitted follows from the game theoretical model of an auction as applied to the use of threat displays. If an agonistic interaction is seen as a bidding process in which the individual who is willing to escalate the most wins, then a threat display that announces how much the signaler is going to bid guarantees losing. A signal that correctly identifies the signaler's state, that is, the intended bid, allows the opponent to out-bid by a small amount every time. In such a situation, it is clearly best to keep one's bid secret. In a 1979 analysis of avian threat display use, Peter Caryl concluded that there was very little evidence to support the traditional view that there was bid-like information in the signal. Few, if any, threat displays were followed by an attack with more than even probability; different threat displays in the same species did not precede attacks with remarkably different probabilities; and a single threat display may 'predict' attack with the same likelihood of predicting abandonment of the contest. In the early 2000s, several authors critically reviewed these conclusions, pointing out that threat displays did consistently predict, albeit with low predictability, subsequent aggressive escalation on the part of the signaler. Caryl's conclusion that the quality of the information transmitted is poor, remains true. These are, at best, quite ambiguous signals.

Is Communication an Arms Race?

A wider case against honest communication was advanced in a number of highly influential articles by John Krebs and Richard Dawkins in the late 1970s and early 1980s. Dawkins and Krebs argued that the exaggerated signals seen in communication between individuals with conflicting interests were attempts to manipulate receivers into acting against their own self-interests. Such hypnotic signals overwhelmed the receiver's senses, hypnotizing them into acting as the signaler's agent. This manipulation of the receiver's behavior was suggested to evolve in concert with ever increasing sales resistance on the receiver's part in an evolutionary arms race.

In contrast to this view of evolutionarily unstable spiralling co-evolution of manipulation and scepticism is the view that signals actually convey useful information, that is, they are basically honest and their use is an evolutionarily stable strategy. Reconciling the latter view with the criticisms leveled against the classical ethological perspective has been a very active research topic, with game theoretical models playing a prominent role. We can address this problem by considering how different types of signals can be used to convey information. For each of these signal types, there is a different reason that receivers may believe the information communicated is reliable.

Types of Signals

Several distinct types of signal have been proposed to be evolutionarily stable against the corrupting pressure of deception or manipulation.

Handicaps

By far the most influential honest signaling hypothesis is the handicap principle, advanced by Amotz Zahavi in a pair of papers in the mid-1970s. Zahavi's verbal model proposed that signal reliability was maintained by the inherently wasteful costliness of a signal. Signalers advertising a desirous or fearsome ability or state could only do so credibly if they use up some of that ability or state. A signal that was wastefully costly to produce could be afforded only by the most able signalers. This verbal model, likening biological communication to a signal of wealth by means of conspicuous consumption, was met with a great deal of scepticism by theoreticians such as John Maynard Smith.

Zahavi's hypothesis gained widespread support only after a formal game theoretical model by Alan Grafen demonstrated that the idea could work, in principle. Zahavi has gone on to suggest that the mechanism underlying handicapped signaling applies to a far wider class of phenomena, claims that have not been widely embraced by researchers in the fields involved. Grafen's models do continue to have widespread support, but some notable criticisms have been made. For example, Tom Getty has questioned whether the critical assumption in Grafen's model, and others like it, that the costs and benefits of different signals are linearly separable, is justified in the biological cases the model is typically applied to. Linearly separable means that, as far as the signaler is concerned, the sources of signal costs are independent of the benefits that a signal will bring. For example, if a courting male uses a handicapping courtship display, then a more attractive display must be more costly when the male is of lower quality. The traditional assumption is that the benefit of a successful courtship is equal for males of all quality. But it may be that males of higher quality are able to turn a successful courtship into greater reproductive success than lower-quality males. If the costs vary with signaler quality, and the benefits do as well, then the two are not linearly separable. Getty concludes that this critical assumption is not justified, and, moreover, that the whole idea of a handicap principle is distractingly unhelpful, and the metaphor ought to be boycotted in favor of less loaded language. The costs of producing and bearing intense signals may prevent some signalers from using them, thereby maintaining honesty, in a process quite like Grafen's models, but without necessarily conforming to Zahavi's larger view of handicaps. Plausible models of signals without linearly separable costs and benefits may produce evolutionarily stable, honest signals in which the cost of producing more attractive signals is prohibitive to signalers of lower quality (the cost of increasing signal intensity prevents signalers of all qualities from exaggerating) yet the absolute cost paid by signalers is zero. Whether this situation can be described as 'handicapping' is debatable. The cost that prevents dishonest exaggeration looks just like a handicap, but that cost is not actually paid. At some point, debating whether it is a handicap or not is less useful than asking how close this model matches what is seen in real biological signals.

Indices

The handicap continues to be a very influential idea, and while it may be the most talked-about form of signal, it is not the only one. Another form of signal that has a much longer, less controversial history is the index. Indices are signals that are honest by way of physical constraint. A large toad makes a deeper croak than a smaller toad and is constrained to do so because of the physics of sound production. If a larger male toad is more attractive to females, and more intimidating to other males, then a deep croak serves as an index of that desirable, fearsome dimension. Whether a specific signal is best described as a handicap, or an index, may be debatable. Some authors, such as Maynard Smith and Harper, see indices as a more widespread, and important, class of signal than do some others. For example, it may be that the depth of the toads croak is exaggerated to a maximum across the population, and the cost of further overcoming the physical constraint on call frequency prevents any further exaggeration. Empirical models of index use such as the sequential assessment game, a model of escalating threat display use based on indices of size, are better supported by empirical data than are handicaps.

Conventional Signals

A more controversial alternative to indices and handicaps are conventional signals, signals without the wasteful cost of a handicap or the physically constrained honesty of an index. A signal is said to be conventional when the meaning of the signal can, theoretically at least, be exchanged with that of another signal. For example, a human in a bar could extend their middle finger upwards and show the back of the hand to another human. This is called 'giving the finger' and involves no more inherent, wasteful, cost than would a similar display using an upraised thumb. One can easily imagine a culture in which a 'thumbs-up' and 'the finger' have their meanings reversed. If the costs and benefits of these signals can be reversed, then their meaning is established by convention; they are conventional signals. Several authors, Zahavi included, have dismissed conventional signals as impossibilities. Game theoretical models of conventional signal use show that they work, at least in principle. The cost imposed by receivers acting on their conventional interpretation of a bluffed signal of strength or desperation in a threat display game can make cheating far less successful than honesty. Some signals, such as the threat displays used by birds, seem more like conventional signals than handicaps or indices.

Some signals do incorporate aspects of both handicapping signals and conventional signals. Vulnerability, or interaction, handicaps are signals that have no inherent cost of production but may bear costs that depend entirely on the receiver's response to the signal, as in conventional signals. Interaction handicaps share with the more classic handicaps described earlier, the property that the form of the signal influences the cost. No cost is paid to produce the signal per se; instead, a cost is paid only if the receiver makes a specific response. The best example of such a signal is a threat display the form of which makes the individual performing the display vulnerable to counterattack from the receiver. The classic example of a vulnerability handicap is the use of a threat display by a newly molted arthropod. Waving their large claw functions as a threat display but entails a great risk of cost, since the claw is effectively useless and vulnerable to damage if (and only if) the receiver reacts to the threat with a counterattack. Although these signals do blur the distinction between conventional signal and handicaps, it is important to realize that conventional signals and handicaps do not grade into one another as general classes of signals.

The costs which maintain honesty in handicaps, if imposed on a game theoretical model in which conventional signals are evolutionarily stable and honest, produce counterintuitive results. Simply adding an arbitrary cost to a conventional signal will not make it extra-resistant to dishonest use. In a model of conventional threat display, if one of the conventional signals is made to be costly in a handicapping sense, then it will be used by the stronger, not weaker, individual. Both the original and posthandicap versions are evolutionarily stable and produce honest signaling, but signalers gain higher fitness in the former, in which signaler quality could be associated with either signal. In the latter, the lower-quality signalers use the handicap and they gain lower payoffs, while everyone else's payoffs remain unchanged. This handicap-enhanced outcome seems unstable in the long run, in that the handicapped signal will not be used at all if a new, costless, signal is made available for use to replace it. The handicap will be used only if the number of signals that can be used is so small that signalers have no option but to include the costly one in their repertoire.

All the three of these signal types: handicaps, indices, and conventional signals, can maintain honesty between individuals with conflicting interests, at least in theory. All have some degree of support from empirical studies. Most likely, all the three do function in stabilizing communication between animals, to some degree, in some cases. However, the debate over the relative importance of these three signal types is far from settled.

Signaling in Biological Contexts: The Degree of Conflict

While animals signal to each other in a wide variety of situations, there are several specific communication scenarios that have interested researchers. Begging signals, courtship displays, threat displays and signals about, or directed to, predators are some of the best-studied examples of biological signaling. These examples may all be placed on a continuum of conflicting interests, as has been done by William Searcy and Stephen Nowicki. Searcy and Nowicki classify the interests of signalers and receivers as identical, overlapping, divergent or opposing, corresponding roughly to interactions within an individual, between kin, between potential mates, and between individuals in aggressive interactions, respectively. The idea that the form of signals used in a social interaction is influenced by the degree of conflict between the participants traces back at least as far as Krebs and Dawkins. The assumption has been that more escalated, or costly, signals will be used in those interactions in which there is a greater degree of conflict. While the idea makes a great deal of intuitive sense, a number of counterexamples can be made. Aposematic signals, in which a toxic signaler and a potential predator have little conflict, show considerable signal exaggeration nonetheless. Conventional signals, where the cost is least, seem to be most likely used in aggressive interactions where the conflicting interests seem rather more stark. In fact, it may be that animals involved in an aggressive interaction have interests that are far more in common than it seems on first inspection. While both individuals prefer that they prevail and the opponent concede, they are united in preferring not to have a potentially injurious escalated physical battle. This sort of pattern of common interest across some possible outcomes, and conflicting interests across others, not only drives conventional signaling models, but may also make it hard to predict large-scale patterns of signal property from the overall degree of conflict.

Honesty and Deception, or Ambiguity?

Various game theoretical models show that signaling can be honest in the face of conflicting interests between signaler and receiver. Empirical studies show that animals do communicate to receivers information that they then put to use. But this does not mean that signals are necessarily honest. Communication between animals seems remarkably ambiguous and imprecise. With the possible exception of alarm calls, where one short call may be used to provide warning to the audience, biological signals seem nothing like the maximally informative, 'say it once, and all is revealed,' traits predicted by most game theory models, or the acme of signals the classical ethologists suggested that evolution selects for. The imprecision of signals may simply reflect some external constraint acting on signals that prevents evolution from removing a large 'noise' component from the signal. According to this view, signalers would benefit if the signals could be made more precise and informative.

An alternative view is that these noisy signals are not failed attempts at maximum honesty, but best thought of as honest-on-average. Alan Grafen and Rufus Johnstone modeled an evolutionarily stable blend of honest and deceptive signalers in a handicapped begging game. The idea that signalers may sporadically exploit receivers while the proportion of honest signals remains high enough to maintain the response in the receivers makes intuitive sense and is widely accepted. The Grafen and Johnstone model, like that of stomatopod threat displays by Eldridge Adams and Michael Mesterton-Gibbons, produces deceptive signaling within a larger population of honest signalers by combining the honest and dishonest signalers together through the use of the same signal. The receiver has to treat the class of all signalers using the same signal as the average of both honest and dishonest types using that signal, in effect never being able to adjust to the types differentially. It is as if the receiver were in a city where 5% of all the $20 bills were counterfeit, choosing to be paid in $20 bills from various sources, but always receiving $19 whenever spending any $20 bill. The receiver chooses to accept $20 bills, without fear of getting stuck with a counterfeit, but never expecting the bill to be worth anything more than $19. Thus, honest and dishonest $20 bills coexist, but the receiver does not have to be described as being 'deceived.'

An alternative to the view that signals are either 'honest' or 'deceptive' is taken by Tom Getty and Peter Hurd, among others. This view proposes that signals might be better characterized as simply 'ambiguous' but still informative. Consider the case of a sentry giving an alarm call. Empirical studies of alarm calling in birds have found examples in which about half the alarm calls are given in the absence of predators, but provide the sentry with the opportunity to take food uncovered by the flock while they take cover from the predator that isn't there. This signal clearly meets the definition of a deceptive signal (below) when given in the absence of a predator. Whenever the alarm is raised, the receivers may either be deceived (by inferring that they were not aware of the possibility that the 'alarm call' signal might be a false alarm), or they may have the correct expectation that a predator has a 50% probability of being present. In the latter case, the receivers may not be deceived, so much as gambling with known odds. This same logic may be extended more productively to situations with more variation. Imagine that a territorial male is either more or less likely to attack an intruder because of some variation in subjective resource value, such as whether a female nesting on his territory is fertile or not. The male has a repertoire of several different threat displays. In each of these different levels of resource value, the male has different probabilities of attacking after each of the different threat displays. If evolution has led to a stable pair of signaler and receiver strategies, then the receiver must be working with an accurate expectation of these probabilities. Outcomes may be better or worse, but the signaler plays the game knowing the odds and cannot properly be described as cheating. If one accepts that evolution has led to an evolutionarily stable pairing of signaler and receiver states, then any signaling system will be honest in this sense.

Definitions

A word about definitions. Perhaps more than any other area within animal behavior, the study of communication has a long history of making liberal use of important terms, with poor, or multiple conflicting, definitions. This may indicate that the central concepts are so intuitively clear that reasonable progress can be made without universally agreed upon formal definitions or that the concepts are so hoary and vague that ideas appearing to be simple require redefining at each use, resulting in a pile of incommensurate crosstalk obstructing the resolution of any of the central issues. The following definitions have been paraphrased from Searcy and Nowicki, and may be subject to all the criticisms mentioned earlier, but still function adequately.

Signal: A signal is a character or behavior that has evolved so as to provide information to other organisms. Signals are usually defined so as to exclude traits that convey information to the detriment of the signaler, such as the rustling noise of a mouse in the grass which is heard by a nearby owl. Some of the more extreme applications of the handicap idea may include such apparently detrimental signals. For example, it could be argued that such life-threatening grass rustling serves as a signal of his quality because of the handicap it imposes on him.

Honesty: For the purpose of this article, honesty will be usually mean 'reliable,' that is to say: there is some variable state that the signaler is in, or aware of, that the receiver cannot know directly. The receiver would benefit from knowing this state. The signal chosen by the signaler allows the receiver to choose a response that is appropriate for the actual state.

Deception: A signal X is said to be deceptive if it elicits a response Y from the receiver which benefits the signaler, and the response Y would be appropriate, beneficial to the receiver, if the state that the signaler is in, or aware of, were a different one from the actual state.

See also: Agonistic Signals; Alarm Calls in Birds and Mammals; Game Theory; Mating Signals; Parent–Offspring Signaling.

Further Reading

Adams ES and Mesterton-Gibbons M (1995) The cost of threat displays and the stability of deceptive communication. *Journal of Mathematical Biology* 175: 405–421.

Bradbury JW and Vehrencamp SL (1998) *Principles of Animal Communication.* Sunderland, MA: Sinauer.

Caryl PG (1979) Communication by agonistic displays: What can games theory contribute? *Behaviour* 68: 136–169.

Dawkins R and Krebs JR (1978) Animal signals: Information or manipulation? In: Krebs JR and Davies NB (eds.) *Behavioural Ecology: An Evolutionary Approach,* 1st edn., pp. 282–309. Oxford: Blackwell.

Getty T (1997) Deception: The correct path to enlightenment? *Trends in Ecology and Evolution* 12: 159–160.

Getty T (2006) Sexually selected signals are not similar to sports handicaps. *Trends in Ecology and Evolution* 21: 83–88.

Hurd PL and Enquist M (2005) A strategic taxonomy of biological communication. *Animal Behaviour* 70: 1155–1170.

Johnstone RA and Grafen A (1993) Dishonesty and the handicap principle. *Animal Behaviour* 46: 759–764.

Krebs JR and Dawkins R (1984) Animal signals: Mind-reading and manipulation. In: Krebs JR and Davies NB (eds.) *Behavioural Ecology: An Evolutionary Approach,* 2nd edn., pp. 380–402. New York, NY: Sinauer.

Maynard Smith J and Harper DGC (2003) *Animal Signals.* Oxford: Oxford University Press.

Searcy WA and Nowicki S (2005) *The Evolution of Animal Communication: Reliability and Deception in Signaling Systems.* Princeton, NJ: Princeton University Press.

Zahavi A (1975) Mate selection: A selection for a handicap. *Journal of Theoretical Biology* 53: 205–214.

Zahavi A (1977) The cost of honesty (further remarks on the handicap principle). *Journal of Theoretical Biology* 67: 603–605.