

Costly signalling theories: beyond the handicap principle

Ben Fraser

Received: 19 July 2011 / Accepted: 28 October 2011 / Published online: 8 November 2011
© Springer Science+Business Media B.V. 2011

Abstract Two recent overviews of costly signalling theory—Maynard-Smith and Harper (2003) and Searcy and Nowicki (2005)—both refuse to count signals kept honest by punishment of dishonesty, as costly signals, because (1) honest signals must be costly in cases of costly signalling, and (2) punishment of dishonesty itself requires explanation. I argue that both pairs of researchers are mistaken: (2) is not a reason to discount signals kept honest by punishment of dishonesty as cases of costly signalling, and (1) betrays too narrow a focus on certain versions of costly signalling theory. In the course of so arguing, I propose a new schema for classifying signal costs, which suggests productive research questions for future conceptual and empirical work on costly signalling.

Keywords Costly signal · Handicap principle · Zahavi, A. · Punishment

Introduction

The problem of reliability in signalling theory consists in explaining why signallers typically do not deceive receivers, given the incentive to do so. One solution to the problem appeals to the costs of signalling to explain reliability. According to costly signalling theory, signals can be reliable even when sender and receiver interests conflict if deception, while possible, is not profitable. Costly signalling theory has gone through several incarnations, some more popular than others. Here, I discuss two recent overviews of costly signalling theory: Maynard-Smith and Harper (2003) and Searcy and Nowicki (2005). Both pairs of researchers refuse to count signals

B. Fraser (✉)

Department of Philosophy, School of Philosophical, Historical and International Studies,
Monash University, Building 11, Melbourne, VIC 3800, Australia
e-mail: ben.fraser@monash.edu

that are kept honest by costly punishment of dishonesty, as costly signals. I argue that this is a mistake.

After briefly sketching the problem of reliability, I will describe several versions of costly signalling theory. Next, I will present Maynard-Smith and Harper and Searcy and Nowicki's denial that signals kept honest by punishment of dishonesty count as costly signals. I will point out a tension between this denial and other claims these researchers make about costly signalling. Then, I will respond to the two key claims underlying the denial: (1) that honest signals must be costly in cases of costly signalling, and (2) that costly punishment of dishonesty itself requires explanation.

I argue that (2) is not a reason to discount signals kept honest by punishment of dishonesty as cases of costly signalling, and that (1) betrays too narrow a focus on certain versions of costly signalling theory, to the neglect of the crucial insight underlying this solution to the problem of reliability. In the course of so arguing, I propose a new schema for classifying signal costs, which suggests productive research questions for future conceptual and empirical work on costly signalling.

The problem of reliability

Organisms often find occasion to be interested in unobvious qualities of other organisms. Rivals for resources—food, territory, mates—want to know each others' fighting ability. Parents want to know how hungry their offspring are. Predators literally on the prowl want to know the evasive ability of potential prey. Organisms metaphorically on the prowl want to know the quality of potential mates. Adaptive decision-making—choosing who to fight, who to flee from, who to feed on and who to mate with—depends crucially on estimating these unobvious qualities.

Often, potential partners, predators and prey themselves provide relevant information: they roar, beg, stot, sing or dance, for example. In short, they signal. The problem of reliability arises when we ask why signal senders do not mislead signal receivers, given the often strong incentive to do so. Surely, signallers should present themselves as stronger, faster, hungrier, or more desirable than in fact they are, if doing so is to their advantage.

The problem of reliability has been, and continues to be, “central to theories of the evolution of signals” (Maynard-Smith and Harper 2003: 6). The problem arises in many contexts, including mate choice, agonistic interactions between conspecifics, alarm calling, and predator–prey interactions.

Of course, there is not always an incentive to deceive in biological signalling systems. Signaller and receiver interests sometimes coincide. For example, a lost joey has an interest in being found, and its mother an interest in finding it. There is no profit for the joey in giving a ventriloquistic distress call. Thus, one answer to the question ‘why are signals reliable?’ is ‘because there is *no* incentive to deceive’. Reliability can be maintained in some signalling systems because signallers and receivers have *coincident interests*.

Also, deception is not always possible. In some cases, reliability may be maintained because signals are *constrained* to be honest: an individual without the

relevant quality simply cannot give the signal. Signals in such cases are termed “indices” (Maynard-Smith and Harper 2003: 15; Searcy and Nowicki 2005: 16–17), and will receive more attention toward the end of this paper (see section “[Beyond the handicap principle](#)”).

However, signaller and receiver interests conflict often enough, and in cases where signalling dishonestly seems at least possible, that there is a real puzzle concerning the persistence of some signalling systems. Under such conditions, shouldn't signalling systems swiftly descend into a cacophony of dishonest proclamations, and eventually fall silent entirely? If everybody lies, nobody will listen, and if nobody listens, nobody will bother lying.

Costly signalling theories

Amotz Zahavi's solution to the problem of reliability was to suggest that signal reliability could be maintained, even in the face of incentives to deceive, if signals of a particular quality were too costly for individuals lacking that quality to produce. In this first formulation of Zahavi's “handicap principle” (1975), costs supposedly keep signals of quality honest by imposing a test on signallers that only genuinely high-quality individuals can pass. Zahavi illustrated his proposal with examples of sexually-selected traits.

An individual with a well developed sexually selected character is an individual which has survived a test. A female which could discriminate between a male possessing a sexually selected character, from one without it, can discriminate between a male which has passed a test and one which has not been tested. Females which selected males with the most developed characters can be sure that they have selected from among the best genotypes of the male population (Zahavi 1975: 207).

For example, according to Zahavi's initial version of costly signalling theory, an extravagant and cumbersome tail is a reliable signal of high quality because only a high quality peacock could afford the costs of developing and displaying it; a lesser peacock so encumbered would succumb to starvation or predation. Signalling in this version of costly signalling theory is on–off, and signal costs are uniformly high.

The problem with Zahavi's initial statement of the costly signalling idea is that signalling in the natural world seems, simply, not to work this way. It is not the case that all peacocks, for example, display tails of identical size, pattern, and brilliance. The same goes for signallers of other species. Signals vary across individuals, in size, intensity, frequency, or along some other dimension. Signalling is not on–off, and not all signals are equally costly.

This realization led Zahavi to reformulate the handicap principle. In his later work, he acknowledged that in costly signalling, “the phenotypic manifestation of the handicap is adjusted to correlate to the phenotypic quality of the individual” (1977: 603). The crucial difference between the initial and the revised versions of the costly signalling idea is as follows.

In the initial version of costly signalling theory, high quality and low quality individuals were assumed to pay the costs of signalling and it was assumed that only high quality individuals survived (at least for very long). In the later version, high signalling costs are paid only by those who can afford those costs, that is, the genuinely high quality individuals. Individuals who cannot afford high signalling costs either do not signal at all, or they signal at a lower intensity that is affordable given their quality. Signals in former case have been termed “conditional” handicaps, and in the latter, “revealing” handicaps (Searcy and Nowicki 2005: 10).

Despite its initial incarnation receiving a sceptical welcome (e.g. Maynard-Smith 1976), the later versions of the costly signalling idea eventually gained acceptance after mathematical modelling work by Grafen (1990a, b) and Johnstone (1995) showed their feasibility in principle. It is now generally agreed that signal costs are one way of solving the problem of reliability.

Costly punishment and reliability

There has been a great proliferation of work on costly signalling since Zahavi introduced the idea, both conceptual (refining the theory) and empirical (describing cases of costly signalling). Two overviews of costly signalling theory—Maynard-Smith and Harper (2003) and Searcy and Nowicki (2005)—have explicitly aimed to bring some order to the field. An important part of the project in both Maynard-Smith and Harper and Searcy and Nowicki’s overviews was the classification of different kinds of signals. Notably, both pairs of researchers *refused* to count as costly signals those signals that are kept reliable by punishment of dishonesty, even though such punishment inflicts costs on dishonest signallers. Maynard-Smith and Harper wrote:

There is one context in which we prefer not to call a signal a handicap, although it may have costly consequences. Suppose there is ‘punishment’. An individual indicates that it will do A, but then does B, and is punished by the receiver: lying signals are costly. Our reason for not referring to such cases as handicaps is that at an evolutionary equilibrium, lying will be rare or absent, and signals can be both cost-free and reliable. The problem lies, not in explaining reliability, given punishment, but in explaining why the receiver is willing to incur the cost of punishing (Maynard-Smith and Harper 2003: 8).

A potential terminological confusion must be cleared up at this point. By “handicap,” Maynard-Smith and Harper mean “a signal whose reliability is ensured because its cost is greater than required by efficacy requirements; the signal may be costly to produce, or have costly consequences” (2003: 15). What Maynard-Smith and Harper call handicaps, I am here calling costly signals. My choice of terminology is motivated by the desire to avoid identifying the costly signalling solution to the problem of reliability too closely with Zahavi’s initial presentation of the handicap principle. As the previous section suggested, and as will be further argued below (see section “[Cost and honesty](#)”), the costly signalling idea covers more than just the idea that high signal costs can winnow out low-quality pretenders by imposing a survival test on signallers.

Like Maynard-Smith and Harper, Searcy and Nowicki refuse to count signals kept reliable by costly punishment of dishonesty, as costly signals. They consider this idea to be:

so far from the original spirit of the handicap principle, in which signals produce costs because of their extravagance, that [this] should be regarded as a separate hypothesis (Searcy and Nowicki 2005: 217).

There are again some potential confusions that must be forestalled here. Searcy and Nowicki subsume costly punishment of dishonesty under the more general heading of “individually directed scepticism,” which can occur when “receivers can recognise the signals of individual signallers, remember the reliability of those individuals in past interactions, and adjust their present response in light of that past experience” (2005: 73). In one form of individually directed scepticism, “receivers refuse to respond in the future to the signals of individuals that have deceived them in the past” (2005: 73). In another form, “receivers retaliate directly, that is with aggression, against signallers they recognize as being consistently deceptive” (2005: 73). Both forms of individually directed scepticism impose a cost on deceptive signallers. Regarding the second form, though, Searcy and Nowicki write, “we know of no case in which this kind of receiver-dependent cost of deception has been documented” (2005: 75). And, although they do allow that costly punishment of dishonesty is a theoretically possible way of maintaining reliability in a signalling system, they insist that this would be a different solution to the problem of reliability than the one provided by the handicap principle, that is, the idea that signals can be kept reliable by their costs (see e.g. Searcy and Nowicki 2005: 9, 11, 179).

To clarify my disagreement with Searcy and Nowicki, the first thing to note is that my point here is primarily a conceptual one, about which kinds of signals are properly counted as costly signals. My contention (argued for at length in what follows) is that signals kept reliable by costly punishment of dishonesty are properly counted as costly signals. This point is independent of the empirical question of whether, or how often, such a solution to the problem of reliability is implemented in actual signalling systems. On that question, though, it should be noted that Searcy and Nowicki are primarily concerned with signalling among non-humans, and when they note the lack of actual cases of costly punishment of dishonesty, it is non-human organisms they have in mind. It may be that costly punishment of dishonesty occurs overwhelmingly (if not solely) in human signalling systems. But this would not vitiate the clarificatory conceptual point I seek to make here. And, it would make that point even more important when it comes to understanding reliability in human signalling systems.

To sum up, it seems that there are two reasons being offered by Maynard-Smith and Harper and Searcy and Nowicki for denying that signals that are kept honest by the punishment of dishonesty count as costly signals. These reasons are:

1. Honest signals in such signalling systems need not be costly.
2. Costly punishment itself requires explanation.

At this point, one may well wonder why the classification of this certain class of signals as costly signals, or otherwise, matters. The issue matters for at least two

reasons. I will present these in general terms here, to set up the more detailed discussions in later sections of the paper.

First of all, considering the issue reveals some internal tensions in the view of costly signalling put forward by Maynard-Smith and Harper and Searcy and Nowicki. Insofar as these two overviews are significant works in the field, and moreover works that explicitly aim to clarify costly signalling theory, it is important to identify and resolve these tensions.

Secondly, close attention to the reasons given for refusing to count such signals as costly signals reveals that both pairs of researchers are overly focussed on certain versions of the costly signalling idea. This focus obscures the key insight behind costly signalling theory. Refocussing on that insight is worthwhile, since it brings to the fore questions that can usefully inform future research on costly signalling.

Below, I elaborate upon and defend these general responses to Maynard-Smith and Harper and Searcy and Nowicki's refusal to count signals that are kept reliable by punishment of dishonesty as costly signals.

Counting costs

The notion of 'cost' is perhaps used a little too freely in discussions of costly signalling. A better term would probably be 'expenditure', where this includes using up energy or resources or time, as well as incurring certain kinds of risk. It is crucial to costly signalling models that expenditures decrease some component of fitness, but just how expenditures link up with fitness costs is a complex matter (Kotiaho 2001). When I use the term 'cost' below, it should be understood as shorthand for 'expenditure that decreases some component of fitness'.

There is some reason to think that Maynard-Smith and Harper and Searcy and Nowicki ought, by their own lights, to count signals kept reliable by punishment of dishonesty as costly signals. To see why this is, first consider the way in which Searcy and Nowicki classify the kinds of costs that can function to underwrite reliability in signalling systems. Their classification system—which draws substantially on that of Sandra Vehrencamp (2000: 279)—can be represented in Table 1.

Receiver-dependent costs “stem from some response of the receiver” (2005: 14). Such costs can come in the form of the risk that the signaller will be attacked by the receiver (a “receiver retaliation” cost), or, of the risk that the signaller will be injured should the receiver decide to attack (a “vulnerability” cost).

Table 1 Classification of signal costs, following Searcy and Nowicki (2005)

	Developmental	Production	Maintenance
Receiver-dependent		Vulnerability Retaliation	
Receiver-independent	Energy Resources	Energy Time Predation	Energy Foraging difficulty Predation

Receiver-independent costs are “imposed regardless of whether or how receivers respond” (2005: 15). These costs are further divided into developmental, production, and maintenance costs. Developmental costs are paid during the time the signal develops. They can include the resources needed to build display structures as well as those needed to build the capacities underlying display behaviours. Production costs are paid when the signal is sent to the receiver. They can include the energy required to send the signal, the time spent sending the signal, and the predation risk that sending the signal entails. Maintenance costs are ongoing costs, the consequences of having to live with a display structure. They can include increased energy expenditure, increased predation risk, and decreased foraging success.

The inclusion of receiver-dependent production costs in their classification scheme shows that Searcy and Nowicki are happy to count the costly consequences of signalling as one of the kinds of signal cost that can underwrite reliability. Likewise, Maynard-Smith and Harper are not reluctant to include costly consequences *per se* in their account of costly signalling. For instance, they count fulmars’ breast-to-breast display during food contests as a costly signal (2003: 7–8; see also Enquist et al. 1985). The relevant cost in this case is the heightened risk of injury if attacked while in this especially vulnerable pose. This is a receiver-dependent production cost, in Searcy and Nowicki’s terms.

Maynard-Smith and Harper also count predation risk as one of the kinds of costs that can underwrite reliability in signalling systems (see e.g. 2003: 17, where they discuss the costs to peacocks of bearing such extravagant tails). It is worth highlighting the fact that predators can occupy two different roles in signalling interactions.¹ One role is that of signal receiver. This is the role occupied by predators in the case, for example, of stotting (see Caro 1986a, b). Another role predators can occupy is that of a third party: an observer or eavesdropper. In this kind of case, the signal is not directed at the predator, but is potentially noticed by the predator, whose response is an important factor in understanding why the signal is reliable. For example, in many birds, begging calls are directed at parents by chicks, and the risk of predators overhearing such calls and plundering the nest is thought to be one of the costs (perhaps the most significant) keeping such signals honest (for more on this, see the discussion and references in section “[Beyond the handicap principle](#)” below).

To return to the main line of argument, here is the tension between Maynard-Smith and Harper and Searcy and Nowicki’s view of signal costs, on the one hand, and on the other, their denial that signals kept reliable by punishment count as costly signals: the costs inflicted on dishonest signallers by punishers look like just another sort of receiver-dependent production cost. Thus, if Maynard-Smith and Harper and Searcy and Nowicki include signals kept honest by receiver retaliation, vulnerability, and/or predation risk as costly signals, then it seems they ought to also count as costly signals those kept honest by punishment of dishonesty. Consistency requires including the signals kept honest by punishment, or excluding those kept honest by, for instance, predation risk. But, signals kept honest by predation risk are among the paradigm cases of costly signalling (again, Maynard-Smith and

¹ Thanks to an anonymous referee for urging me to emphasise this point.

Harper think this is an important part of the explanation for why peacocks' tails are reliable signals). Thus, the consistency requirement seems to exert pressure in one direction in particular, that is, to include the signals kept honest by punishment of dishonesty as costly signals.

Of course, Maynard-Smith and Harper and Searcy and Nowicki will no doubt insist that predation risk and punishment of dishonesty are relevantly different costs. As the quotes given earlier indicate, their reasons for not counting signals kept honest by punishment of dishonesty as costly signals have to do with the cost of *honest* signals in such signalling systems, and with the need to explain punishment itself. I discuss these issues below, starting with the need to explain why punishers punish.

Explaining costly punishment

Maynard-Smith and Harper rightly note that an explanation for signal reliability that appeals to punishment of dishonesty is importantly incomplete, insofar as such punishment is costly not just to the dishonest signaller, but to the punisher as well. The willingness of receivers and observers of dishonest signals to sometimes take the time, spend the effort, and run the risks of dishing out punishment is a puzzle in its own right.

To better appreciate the puzzle, consider a parallel with the attempt to explain the evolution of cooperation by appeal to punishment of defection. Punishment of free riders—those who take the benefits of cooperation within a group but do not contribute to the production of those goods—can stabilise cooperation in a population (Boyd and Richerson 1992). But, there remains the issue of what stabilises punishment: a so-called 'second order' free rider problem exists. Punishment itself can be seen as a kind of cooperation, insofar as it is costly to the punisher but benefits other members of the group (by weeding out free riders). Punishment of free-riding itself thus stands in need of explanation, and an account of cooperation in terms of such punishment is incomplete without that further explanation.

The same kind of problem exists for punishment as a means of stabilising reliability in a signalling system. Maynard-Smith and Harper are right that, insofar as punishing dishonest signals is costly to the punisher, some further story must be provided about why punishers punish. But while this is a demand that must be met, it is not a reason to disqualify such signals as costly signals.

To see why, notice that in the case of other receiver-dependent costs—predation risk, for instance—the same demand is present. It is not free for a predator to take prey. Doing so involves costs in energy and time, and perhaps some element of risk to the predator itself. Of course, an account of why predators are willing to incur the costs of taking prey is easily provided: such costs are recouped (and more) when the prey is consumed. Similar things could be said about the fulmars' breast-to-breast display during food contests (a display counted as a costly signal by Maynard-Smith and Harper). It is not free for the receiver of this signal to actually attack the posturing bird, since attacking costs time and energy and carries the risk of injury

for the attacked too. But attacking is fairly readily explained by appealing to the benefits to be had, in this case, potentially obtaining the valuable morsel of food being contested.

In the case of punishment, the story is likely to be more complex. Nevertheless, that complexity in itself is not grounds for refusing to count signals kept honest by punishment of dishonesty as costly signals. The crucial point is that, even if punishing does impose some costs on the punisher, punishing may nevertheless also secure benefits for the punisher that outweigh those costs, making punishment something that works, overall, to the punisher's advantage.² The benefits of punishing will most likely vary across cases. One interesting—if still largely speculative—idea is that punishment itself might serve as a costly signal (Smith and Bliege-Bird 2005; Miller 2007).

Speculation aside, though, the bare point stands. Maynard-Smith and Harper and Searcy and Nowicki certainly do not establish that an explanation for why punishers punish *cannot* be given. Even if giving it in any particular case is tricky, the fact that punishment needs to be explained should not be taken as grounds on which to disqualify signals kept honest by punishment of deception from counting as costly signals. To reiterate: in general, receiver-dependent costs require explanation, and this fact alone does not disqualify signals kept reliable by such costs from counting as costly signals (by Maynard-Smith and Harper's and Searcy and Nowicki's own lights).

Cost and honesty

Assuming for simplicity's sake that there is no 'noise' in the signalling system, if reliability is maintained by the punishment of dishonesty, then only dishonest signallers will pay high signal costs; individuals signalling honestly may do so cheaply or even for free. For this reason—honest signals need not be costly—Maynard-Smith and Harper and Searcy and Nowicki balk at counting such signalling systems as cases of costly signalling. But to claim that honest signals must be costly in order for costly signalling to be the right explanation for reliability in a signalling system, is to reveal an overly narrow focus on certain versions of costly signalling theory, and to neglect the key insight underlying this kind of solution to the problem of reliability.

Above, I distinguished between Zahavi handicaps, and conditional and revealing handicaps. In the first kind of costly signalling, signal costs are uniformly high, and thus honest signals are costly. In the second kind of costly signalling, signals costs are high and only honest signallers send signals, so again honest signals are costly. In the third kind of costly signalling, signal display is adjusted to signaller quality, so low quality but honest individuals signal their low quality with a relatively (compared to high quality individuals) low expenditure of energy, resources, or risk.

² And, as Kim Sterelny has pointed out to me, punishment need not impose *any* extra costs on signal receivers. If known liars are punished via ostracism—not being chosen as partners for mutually profitable interactions—then punishment takes place as a side-effect of partner choice, and involves no costs that discriminating individuals were not paying anyway.

Honest signals in this case may involve little expenditure. In the case of revealing handicaps, low quality individuals are not unable to signal at the same level as high quality individuals—reliability is not maintained because dishonesty is impossible—but doing so would be unprofitable for them, since the benefits gained by deceiving others are (the theory supposes) outweighed by the costs of signalling at a higher level than their actual quality can support. Signalling a given level of quality is more costly, on balance, for one who is not in fact of that quality than for one who is.

Here, the key insight behind costly signalling theory becomes apparent: it is not absolute signal costs that are relevant, but *differential* costs. This point has been made previously in the costly signalling literature (see e.g. Grafen 1990a, b; Johnstone 1995; Saunders 2009). It is also a point made by Maynard-Smith and Harper (2003: 19–20) and Searcy and Nowicki (2005: 11–13) themselves. It nevertheless bears repeating here, since appreciating this point is crucial in seeing why signals kept reliable by punishment of dishonesty are properly counted as costly signals.

As in the case of Zahavi handicaps, if perhaps not so obviously, it is differential costs that are crucial in the case of conditional handicaps, where individuals signal only if they actually possess the relevant quality. In this version of the theory, whatever benefits are gained by pretending to possess a quality one lacks, are cancelled out (and more besides) by the costs of signalling dishonestly. Signalling dishonestly, while possible, is not profitable.

Less obviously still, it is differential rather than absolute costs that underwrite reliability even in the case of Zahavi handicaps. There, the absolute cost of signalling—the expenditure of energy, risk, or other resources required to send the signal—is supposed to be the same for all signallers, high and low quality alike. High and not low quality signallers pass the ‘test’ that signalling imposes, because the cost of signalling represents less of the overall ‘budget’ available to high quality individuals for signalling and all other activities, such as foraging, mating, and rearing of offspring. Signal costs are differential here, even though the absolute cost is the same for high and low quality individuals, because the relative cost—the proportion of the individual’s overall budget that signalling costs deplete—is lower for the former than for the latter.

To bring this all together: in many cases of costly signalling the cost of honest signals may be high, but what is crucial to the maintenance of reliability is differential cost. A given signal must be more costly for a dishonest than an honest signaller, and signalling dishonestly must be costly enough that doing so is not, on balance, worth it. Crucially for current purposes, the requirement of differential signal cost can be met *even if honest signals are free*.

To see this, consider a simple model constructed by Lachmann et al. (2001). In this model, signallers vary in some quality of interest to receivers, the higher a signaller’s quality is perceived to be by receivers the better off that signaller is, and receivers benefit from accurately assessing signallers’ quality. Importantly, signalling one’s quality up to its true level is free, but any exaggeration is lethal.

In this model, signalling a certain level of quality is *differentially* costly. This is so even though honest signallers pay *nothing at all* to send that signal. The reason is

that the signal is costly for *dishonest* signallers. Dishonest ‘self-aggrandisers’ pay with their lives. Dishonest ‘self-deprecators’ lose out on the benefits that would come from signalling their true, higher quality. Under these conditions, signallers would signal their true quality, and receivers would trust those signals, not because honest signals are costly but because dishonest ones are. Thus, honest signals need not be costly in order for differential costs to underwrite reliability in a signalling system.

Rather confusingly, Lachmann et al. seem to think their model in some way undermines costly signalling theory, by showing that “contrary to the primary tenet of costly signalling theory, waste is *not* necessary” for honest signals (2001: 13190). I say ‘confusingly’ because Lachmann et al. themselves write:

The critical idea behind Zahavi’s original formulation of the handicap principle is that the necessary incentives come from the costs of signalling. When signalling is costly—and differentially costly according to the meaning and/or validity of the message—signallers can be relied upon to report aspects of their condition or of the environment (2001: 13189–90).

The importance of differential costs is recognized here, but its centrality is perhaps obscured by a focus on Zahavi’s first version of costly signalling theory. The ‘critical idea’ behind Zahavi’s initial formulation of the handicap principle was indeed that signals of high quality can be reliable if such signals are so costly that only genuinely high quality individuals can afford to send them. But even if it was central to Zahavi’s initial view, the idea that honest signals must be high cost is not essential to costly signalling theory *tout court*. Rather, the essential idea is that of differential costs, which—as Lachmann et al’s own model shows—can obtain even when honest signals are free.

Cost-free honesty is entirely compatible with the key idea underlying costly signalling theory. Thus, the fact that honest signals need not be costly in signalling systems where reliability is maintained by the punishment of *dishonesty* does not rule out such signalling systems from counting as cases of costly signalling.

Beyond the handicap principle

Searcy and Nowicki shared Maynard-Smith and Harper’s reservations about counting signals kept honest by punishment of dishonesty as costly signals. They considered the idea to be “far from the original spirit of the handicap principle” (2005: 217). I hope by now it is apparent where this claim goes wrong. In this final section, I want to shift to a more constructive mode. I will return to the issue of counting kinds of costs in costly signalling, and offer a revision of Searcy and Nowicki’s classificatory scheme, and show how this new classification suggests questions for future research on costly signalling.

Recall, Searcy and Nowicki’s major division was between receiver-dependent and receiver-independent costs. In their scheme, the various signal costs resulting from the actions of other organisms are distributed on either side of that divide. Risks posed to the signaller by receivers are receiver-dependent costs, whereas risks

posed to signallers by organisms other than the receiver, such as predators or conspecific rivals, are receiver-independent costs. Given this, it may clarify matters to reorganize Searcy and Nowicki's division of signal costs (see Table 2).

In this scheme, the major divide lies between costs a signaller pays regardless of what receivers or any other organisms do, and costs that depend in some way on the actions of receivers or other organisms. I have called this the distinction between *intrinsic* and *imposed* costs.

Before discussing this proposed classificatory scheme in detail, and drawing out some of the research questions it suggests, it is worth clarifying the relationship between costly signals, thus understood, and the above-mentioned (see section “[The problem of reliability](#)”) indices.

An index, as defined by Maynard-Smith and Harper (2003: 15), is “a signal whose intensity is causally related to the quality being signalled, and which cannot be faked.” It is important to note that I have not folded indices into my classification of costly signals. It is true that some indices have high intrinsic costs. For example, Maynard-Smith and Harper suggest that roaring by red deer might be best understood as an index: depth of roar is a reliable signal of size, since depth of roar depends on vocal tract length, which depends on size, and hence reliability is maintained by constraints against deception.

While the empirical details of red deer roaring remain controversial (as Maynard-Smith and Harper acknowledge), this description will suffice for current purposes. Roaring is energetically demanding, but it is not the energetic demands *per se* that explain the reliability of roaring as a signal of size. Rather, it is the fact that the causal link between the quality signalled and the nature of the signal constrains signallers to be honest.

To fully drive home this point, notice that there are indices that do not have high intrinsic costs. For example, Maynard-Smith and Harper also discuss territorial disputes in funnelweb spiders to illustrate the distinction between handicaps and indices. Spiders confronted by a rival more than 10% bigger than them withdraw without a fight. Spiders signal their size to each other by plucking the web on which the confrontation takes place. The extent to which a spider is able to vibrate the web is dictated by the spider's size; for a petite spider, faking heftiness is not an option. And notice, plucking the web requires only a miniscule expenditure of energy.

So, to be clear, appealing to costs, on the one hand, or to constraints on the other, remain two distinct ways in which to address the problem of reliability.

Table 2 Intrinsic versus imposed costs

	Developmental	Production	Maintenance
Imposed		Vulnerability Retaliation Predation Punishment	Predation
Intrinsic	Energy Resources	Energy Time	Energy Foraging difficulty

One reason to opt for this way of classifying signal costs is simply that it seems natural to group the various imposed costs together. Also, it is incumbent upon one who accepts Searcy and Nowicki's schema (and those like it, such as Vehrencamp's [2000] mentioned above) to give an account of what distinguishes the *receivers* of a signal from what might be called *eavesdroppers*.

This may seem an easily-shouldered burden. When giving their definition of a signal, Maynard-Smith and Harper (2003: 3) say that the effect of a trait (behavioural or morphological) on other organisms' behavior must be part of the explanation for why that trait evolved in order for the trait to count as a signal (as opposed to merely being a "cue", that is, a feature that other organisms use as a guide to action). This suggests that it is the organisms whose responses help explain the adaptive value of a trait to the signaller that should be counted as the receivers of the signal. Organisms whose responses to the signal do not help explain why the signaller evolved that particular behaviour or morphological trait, should be counted as mere eavesdroppers.

However, while this may in principle divide receivers from eavesdroppers, telling the two apart in practice may sometimes be quite difficult. For example, stalking lions are certainly receivers of a stotting gazelle's signalling behavior, but what of the gazelle's conspecifics? Perhaps nearby sexual rivals and potential mates note the gazelle's athleticism and alter their behaviour toward it (e.g. yielding to it in confrontations and accepting its advances, respectively), thus providing additional selective pressure in favor of stotting and qualifying themselves as signal receivers.³ The empirical difficulties here may well be formidable.

Of course, some way of deciding which organisms are the receivers of a given signal will be required, eventually, by all who hope to fully understand that signal. But, by dividing up costs into intrinsic and imposed, one can count costs without having to first settle potentially very tough empirical questions about the selective history of the trait in question.

On a more positive note, dividing up costs into intrinsic and imposed suggests a potentially fruitful question to be investigated by future research on costly signalling. The question is: why in some cases of costly signalling are intrinsic costs apparently of most importance, and in other cases, imposed costs?

Consider the following two cases. In the case of the peacock's tail, a paradigmatic costly signal, a large (perhaps predominant) role is played by intrinsic costs: the energy and matter required to grow the tail itself. By contrast, in the breast-to-breast display among fulmars (Maynard-Smith and Harper's example, mentioned above), the costs of the signal are primarily imposed, consisting in the risk of attack and injury by the receiver. Why in the one instance of costly signalling are intrinsic costs seemingly crucial, while in the other it is instead imposed costs?

Lachmann et al. suggest an answer to this question. They claim that whether signallers use "costly" signals or "conventional" signals (meaning, in my terms, signals with high vs. low *intrinsic* costs) depends on "the relative ability of receivers to verify the integrity of messages":

³ Thanks to Richard Joyce for this point.

[If] the advertised trait ... only manifests itself stochastically and far in the future, [then a] receiver cannot verify the signal within the time frame of her decision, nor can she identify and punish deception should it occur. Consequently, the costs to deception cannot be imposed through social interaction. Instead, the mechanism of signal reliability must involve differential costs intrinsic to signal production... [But if] receivers can verify the signal with relative ease, should they choose to do so, [then] production costs are not needed to ensure reliable signalling (Lachmann et al 2001: 13191).

In general, Lachmann et al predict, for verifiable qualities we should expect to see signals that have low intrinsic costs but high imposed costs for dishonesty, and we should expect to see “costly” (i.e. high intrinsic costs) signals used otherwise (2001: 13191).

This is an insightful suggestion. It nicely accounts for the difference between the case of the peacock’s tail and the fulmar’s breast-to-breast display. In the former, the signal advertises genetic quality, which cannot be easily or immediately verified (only showing up in the survival and reproductive success of offspring). In the latter, the signal indicates willingness to fight over a piece of food, which can be readily tested by competitors. In line with Lachmann et al’s suggestion, when receivers can relatively easily and immediately detect deception (should it be attempted), signal costs are largely imposed, while high intrinsic signal costs are paid when deception would only become apparent well after the dishonest signal has been observed, believed, and acted upon.

There are two kinds of counter-example one might offer against Lachmann et al’s prediction. One might identify cases where signals advertise relatively easily verifiable qualities but involve high intrinsic costs, or, one might find cases where signals advertise hard-to-verify qualities but involve mostly imposed costs. I think both kinds of counter-example can be offered, although, admittedly, much depends on currently contentious empirical details. The following cases suggest, however, that Lachmann et al’s claims about verifiability and signal costs cannot be taken to settle the question of why costly signalling in some cases relies on intrinsic costs, and in other cases, imposed costs.

Stotting is a classic case of a (supposedly) costly signal. It was one of Zahavi’s own examples (1997: 6–7). ‘Stotting’ describes the behaviour of gazelle that make prodigious leaps into the air upon spotting a stalking predator. These leaps seemingly waste valuable energy that could be better spent running away if the predator decides to give chase. On the costly signalling analysis, though, it is the very cost of stotting itself that deters predators from giving chase, since only fast, agile, healthy gazelle could afford to be so wasteful when under threat.

Whether or not the costly signalling analysis of stotting is correct is an open empirical question (see e.g. Caro 1986a, 1986b). *Prima facie*, though, stotting provides a counter-example to Lachmann et al’s prediction that signal costs will be imposed when verification is relatively easy. As in the case of fulmars, the quality being signalled by stotting gazelle—call it evasiveness—could be readily tested (by a predator giving chase). Nevertheless, the intrinsic costs of stotting play an

important role in explaining its reliability as a signal, on the going costly signalling analysis. Thus, in this case, the link between easy verification and imposed costs comes apart.

Begging calls are another classic case of (supposedly) costly signalling. Chicks would benefit from exaggerating their hunger to gain more food from the feeding parent, while parents would do best if they could accurately assess the hunger level of their offspring. There is thus a conflict of interest here, but begging volume is a reliable signal of hunger level nevertheless.

When it comes to a costly signalling analysis of this kind of case, the empirical issues are formidably complex. Even so, indications are that the intrinsic costs of begging may be rather low, and that the major cost of such signalling comes in the form of predation risk. Sated (or nearly so) chicks do not exaggerate their need for food by calling loudly, because the value to them of excess food is outweighed by the danger of being heard, found, and killed by a predator. Dishonesty is possible, but not profitable, because of imposed costs (Maynard-Smith and Harper 2003: 35–37; Searcy and Nowicki 2005: 45–50; although see Mueno-Rueda and Redondo 2011 for more recent evidence for intrinsic costs of begging).

Crucially for current purposes, offspring hunger level is a hard quality for parents to verify within the timeframe of their decision about which chicks to feed, and how much. A parent could let its offspring starve, and see which one dies first, but that hardly seems feasible. Hunger level looks more akin to a male peacock's genetic quality—in terms of verifiability as Lachmann et al. understand it—than to a fulmar's willingness to fight for a morsel of food, or a gazelle's evasive ability. Despite this, however, the relevant costs of begging calls look to be imposed (in this case, by predators). Thus, in this case, the link between difficult verification and intrinsic costs comes apart.

It is worth reiterating that the relevant empirical details in the cases of stotting and begging as costly signals remain controversial. I do not mean to suggest that my brief treatment here suffices to falsify Lachmann et al's prediction about the relationship between verifiability and intrinsic versus imposed signal costs. I have at best gestured at some potential problem cases for that prediction. My goal here isn't to dismiss Lachmann et al's contribution. It is rather to point out the need for more work on the issue.

Conclusion

The key insight underlying costly signalling theory is that signal reliability can be maintained, despite conflicting sender and receiver interests, if dishonesty, while possible, is not profitable. The insight can apply, even if honest signals are free, because costly punishment of dishonesty counts as a kind of imposed signal cost. Thus—and despite the reservations expressed in Maynard-Smith and Harper and Searcy and Nowicki's recent overviews of costly signalling theory—signalling systems in which reliability is maintained by the costly punishment of dishonesty should be counted as cases of costly signalling. Classifying signal costs into intrinsic and imposed also makes vivid a currently under-discussed question: why in some

cases of costly signalling do intrinsic costs underwrite reliability, while in others the work is done by imposed costs? The relative ability of receivers to verify the signal (and hence to impose costs on dishonest signallers) likely has some important part to play in answering this question. However, this cannot be the whole story. I hope my discussion here, especially my division of signal costs into intrinsic and imposed, will be both spur and guide for future work on costly signalling.

References

- Boyd R, Richerson P (1992) Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethol Sociobiol* 13:171–195
- Caro T (1986a) The functions of stotting: a review of the hypotheses. *Anim Behav* 34:649–662
- Caro T (1986b) The functions of stotting in Thompson's gazelle: some tests of the predictions. *Anim Behav* 34:663–684
- Enquist M, Plane E, Roed J (1985) Aggressive communication in fulmars (*Fulmarus glacialis*) competing for food. *Anim Behav* 33:1007–1020
- Grafen A (1990a) Biological signals as handicaps. *J Theor Biol* 144:517–546
- Grafen A (1990b) Sexual selection unhandicapped by the Fisher process. *J Theor Biol* 144:473–516
- Johnstone R (1995) Sexual selection, honest advertisement and the handicap principle. *Biol Rev* 70:1–65
- Kotiaho J (2001) Costs of sexual traits. *Biol Rev* 76:365–376
- Lachmann M, Szamado S, Bergstrom C (2001) Cost and conflict in animal signals and human language. *Proc Natl Acad Sci USA* 98(23):13189–13194
- Maynard-Smith J (1976) Sexual selection and the handicap principle. *J Theor Biol* 57:239–242
- Maynard-Smith J, Harper D (2003) *Animal signals*. Oxford University Press, Oxford
- Miller G (2007) Sexual selection for moral virtues. *Q Rev Biol* 82:97–121
- Mueno-Rueda G, Redondo T (2011) Begging at high level simultaneously impairs growth and immune response in southern shrike (*Lanius meridionalis*) nestlings. *J Evol Biol* 24:1091–1098
- Saunders S (2009) Costly signalling: a work in progress. *Biol Philos* 24(3):405–416
- Searcy W, Nowicki S (2005) *The evolution of animal communication*. Princeton University Press, Princeton
- Smith E, Bliege-Bird R (2005) Costly signalling and cooperative behaviour. In: Gintis H, Bowles S, Boyd R, Fehr E (eds) *Moral sentiments and material interests*. MIT Press, Cambridge
- Vehrencamp S (2000) Handicap, index, and conventional signal elements of bird song. In: Espmark Y, Amundsen T, Rosenqvist G (eds) *Animal signals: signalling and signal design in animal communication*. Tapir Academic Press, Trondheim
- Zahavi A (1975) Mate selection: a selection for handicap. *J Theor Biol* 53:205–214
- Zahavi A (1977) The cost of honest (further remarks on the handicap principle). *J Theor Biol* 67(3):603–605
- Zahavi A, Zahavi A (1997) *The handicap principle: a missing piece of Darwin's puzzle*. Oxford University Press, Oxford