Animal Behaviour 81 (2011) 3–10

Contents lists available at ScienceDirect

Animal Behaviour

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

Review

The cost of honesty and the fallacy of the handicap principle

Szabolcs Számadó*

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

ARTICLE INFO

Article history:
Received 12 January 2010
Initial acceptance 9 March 2010
Final acceptance 23 August 2010
Available online 22 October 2010
MS. number: 10-00030R

Keywords:
cheating
communication
conflict of interest
handicap principle
honest signalling
potential cost
realized cost
strategic cost

Honest signalling assumes a correlation between an observable signal and a nonobservable quality. There are many mutually nonexclusive mechanisms that can achieve such a correlation; however, for a long time the handicap principle has been identified as the main solution to this problem. In short, it claims that signals need to be costly to be honest and that honest signallers have to pay this extra cost at the equilibrium (i.e. signals have to be handicaps). Honesty, however, is not maintained by the realized cost paid by honest signallers at the equilibrium but by the potential cost of cheating. Whether this potential cost implies a realized cost for honest signallers depends on the biological details of the system and thus this cost cannot be predicted a priori without knowledge of these details. Accordingly, depending on these details, signals need not be costly to be honest, even under conflict of interest. In other words, handicapping equilibrium signals are not the only way to create a high potential cost of cheating. Here I first review the theoretical models supporting the above conclusion, and then I list mechanisms that can maintain a high potential cost of cheating without imposing extra realized cost (i.e. a handicap) on honest signallers at the equilibrium. Identifying and describing those constraints or the lack of them that might create a connection between these two types of cost should be a major research agenda.

The honesty of animal communication is a hotly debated topic. Honest signalling assumes a correlation between an observable signal and a nonobservable quality. How this correlation is achieved is the key question of reliable communication and the topic of debate. Recent reviews agree that honesty can be generated by several mechanisms, most notably by shared interest, constraints (i.e. indices) and finally by signal cost (Vehrencamp 2000; Hurd & Enquist 2005; Searcy & Nowicki 2005). While the handicap principle has been criticized on several accounts (Dawkins & Guilford 1991; Hurd 1995, 1997; Getty 1998; Számadó 1999, 2003; Lachmann et al. 2001; Maynard Smith & Harper 2003; see Getty 2006 for a review on the marginal cost versus marginal benefit issue), it is still widely accepted and cited, in its original form, as the main mechanism that maintains the honesty of signals under conflict of interest. According to the most frequently cited formulation, signals have to be costly to be honest, and thus it is this wasteful signal cost paid at the equilibrium, that is, a handicap, that maintains honesty (Zahavi 1975, 1977; Grafen 1990). However, theoretical models have revealed that it is not the realized cost paid by honest signallers at the equilibrium that maintains honesty, but the potential cost of cheating (Hurd 1995; Számadó 1999; Lachmann et al. 2001). As long as this potential cost is high enough, the realized cost for honest signallers can be zero or even negative (Hurd 1995; Számadó 1999; Lachmann et al. 2001). Here I first review these theoretical models and discuss the implication of these results for empirical research and experimental design. Then I review the empirical literature on signal cost and highlight why the available data are not sufficient to evaluate the importance of signal cost in the maintenance of honesty. Finally, I give a short overview of those mechanisms where the potential cost of cheating can be high enough to maintain honesty without imposing an extra realized cost on honest signallers at the equilibrium. These mechanisms include pooling equilibria (Lachmann & Bergstrom 1998), indices (Maynard Smith & Harper 1995), inclusive fitness cost (Johnstone 1999), frequency-dependent selection (Számadó 2000), individual recognition (Tibbetts & Dale 2007), amplifiers (Hasson 1997), proximity risk (Számadó 2008) and punishment of the cheaters (i.e. social cost; Lachmann et al. 2001).

TERMINOLOGY

Before the discussion of the models it is worth introducing the terminology to avoid potential confusion (see Appendix 1 for a glossary of common terms). First, a distinction has to be made between cues and signals (Maynard Smith & Harper 1995, 2003). While both have the potential to influence the behaviour of the receiver, there is a crucial difference, namely that cues did not

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

0003-3472/$38.00
doi:10.1016/j.anbehav.2010.08.022

evolve to fit this function whereas signals evolved because, on average, they increase the fitness of the signaller by altering the behaviour of the receiver in a favourable way. For example, size can be used as a cue between rivals in a competitive situation, yet size in itself did not evolve to deter rivals. On the other hand, presentation of weaponry in a ready-to-fight form for longer than is necessary for pure fighting efficacy would demand it is a signal (for example 'medial horn presentation' in ungulates, Walther 1984); it evolved exactly because it has the potential to alter the behaviour of rivals in a beneficial way from the point of view of the signaller (i.e. it might persuade them to give up the contest). With regard to signal cost, Maynard Smith & Harper (1995, 2003) distinguished between (1) efficacy cost, a cost necessary to transmit the given information unambiguously, and (2) strategic cost, a cost paid at the equilibrium, on top of the efficacy cost, responsible for the reliability of the signal. Production cost is the sum of efficacy and strategic costs (see Fig. 1). Accordingly, signals can be classified into the following three categories: (1) handicaps, which have both efficacy and strategic costs; (2) minimal-cost signals, which have only efficacy cost; and (3) cost-free signals, which do not have strategic cost and the efficacy cost component is negligible (see Fig. 1). Note that the cost of minimal-cost signals can actually be very high depending on the situation. For example, it is expensive to build a lighthouse even if it does not have any ornamentation in excess. Obviously, in the strict sense every signal has some production cost that comes from the efficacy component even if it is negligible compared to other costs and benefits in the system. This is well accepted and not debated. The source of the debate is whether or not the production cost of signals has to include a strategic cost component to maintain honesty under conflict of interest (as claimed by Zahavi 1975 and Grafen 1990).

Finally, there is another type of cost, which often comes up in the literature related to handicaps: the so-called 'receiver-imposed cost' (Dawkins & Guilford 1991; Vehrencamp 2000; Hurd & Enquist 2005; Searcy & Nowicki 2005). Unfortunately, the source and the nature of this cost are not well defined and there is only one formal approach (Hurd & Enquist 2005). Part of the difficulty is that a potential cost independent of the signalling game is not recognized, so much so that there is no accepted terminology to name it. Of course, considering a cost independent of the signalling game may sound contradictory. Yet it is not; simply because all signalling games are built on a decision problem that exists regardless of signalling. These decision problems are the heart and soul and the ultimate source of all signalling situations. Parents returning to the nest have to decide which offspring to feed, males wishing to reproduce have to decide whether or not to compete for a given territory, females wishing to reproduce have to decide which male to choose. These decisions have to be made regardless of whether there is a signalling game involved or not. A bad decision can be

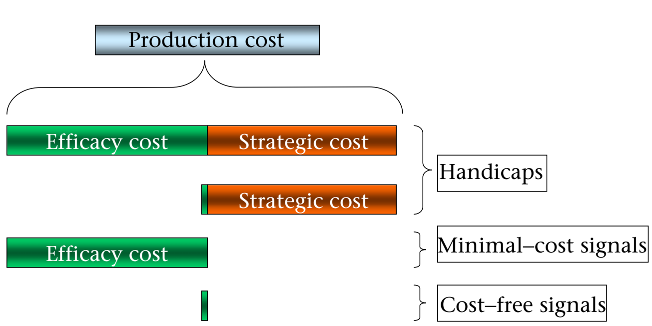

Figure 1. Signal types as a function of production cost. Handicaps have strategic cost, minimal-cost signals have only efficacy cost and finally cost-free signals have neither (i.e. the efficacy cost is negligible).

made even if there is no signalling game at all. What signalling can do is to provide extra information to the animal in that decision situation, which may or may not help to improve the decision, but the cost of making a bad decision will be the same (note that the same argument holds for the use of cues as well). For example, a weak individual fighting against a strong one (as a result of a bad decision) will be beaten up in the same way regardless of whether or not there was a signalling stage before the fight. A signalling stage only helps weak individuals to avoid this bad decision, but it will not alter the cost of the bad decision once it was made. In this light, I call the cost that is independent of the signalling game and that results from making a bad decision in the underlying optimization problem an 'optimization cost'. Potential confusion arises from the fact that this optimization cost need not always be apparent. For example, in the most frequently investigated signalling model, the so-called 'action–response games' (Hurd 1995; Hurd & Enquist 2005), it is always worth asking for more if there is a conflict of interest and the signaller and the receiver are not related; even if they are related, the cost is still not apparent because for the signaller it is always worth asking for more than the receiver is willing to give (yet in this latter situation it is still possible for the signaller to ask for 'too much'; i.e. there is an optimization cost associated with signals asking for more than the signaller's optimum, whereas when the signaller and receiver are not related, there is no optimization cost associated with asking for more, simply because the signaller's optimum is to ask for everything).

GAME THEORY: HONEST SIGNALLING, SIGNAL COST AND EQUILIBRIUM COST

Signalling models used to investigate the honesty of animal communication can be classified into two groups. In all signalling games there is a signaller and a receiver, the receiver has to make a decision that may have an impact on the fitness of both players, and the signaller has some private information (potentially useful for the receiver) which the signaller may or may not reveal by sending signals. Differences between the models result from the assumptions about the possession of the resource at the start of the game and about the range of actions available to the players (these actions are described as the signaller and receiver strategy, respectively). In the first group we have the so-called 'action–response games' (reviewed in Hurd & Enquist 2005). In these games the receiver has the control of the resource at the start of the game and the signaller may or may not send a signal to request the resource. The resource can be divisible or not: in the first case we have a continuous model and in the second a discrete model. The signal and the receiver's response too can either be a discrete state ('ask' or 'not ask', and 'give' or 'not give', respectively) or can change on a continuous scale from zero to one. In the second group we have the so-called 'symmetric' games (or 'mutual signalling games', Hurd & Enquist 2005), where the resource is not in the possession of either player, and where both players have the same strategic options (and both can play the role of the signaller and the receiver). These games are usually used to describe contest situations where players have to fight for the possession of the resource. Most of these games are based on the famous hawk–dove game (Maynard Smith 1974, 1982), and in most of these models signals are discrete.

Grafen's (1990) model, from which he derived his famous 'main handicap results', is a continuous action–response game investigating separating equilibria. At a separating equilibrium, each signaller type gives a signal of different intensity; thus receivers can reliably infer the quality of the signaller from the signal. In contrast to this, there are the so-called pooling equilibria (Lachmann &

Bergstrom 1998) in which signallers of different types can give the same signal. Since pooling equilibria are not entirely honest (by definition), researchers first focused their attention on (fully honest) separating equilibria. Grafen (1990) found that the equilibrium of his model can be described by the following conditions: (1) signalling is honest, that is, signals reveal the true quality of the signaller (this follows from the fact that Grafen investigated separating equilibria), (2) signals are costly (at the equilibrium, i.e. they have a realized cost for honest signallers), and (3) signals are costlier for worse signallers (i.e. signallers in worse condition or with worse genes have to pay more for the signal of the same intensity). Based on these results, Grafen (1990, page 532) drew the conclusion: 'Persuasive signalling necessarily involves waste as only cost can enforce honesty'. This statement, along with the results, was taken as confirmation of Zahavi's handicap theory, which in turn led to the repeated claims of Zahavi (Zahavi & Zahavi 1997, page 230) that 'signals are selected because they handicap the organism in a way that guarantees that the signal is reliable' or that 'The handicap principle is an essential component in all signals' (Zahavi 2008, page 2). Note that the Zahavi–Grafen claim is not that signals have efficacy cost (hence some sort of production cost, since every feature has it) or that signals are subject to cost–benefit constraints (since this too holds for any feature and thus any theory proposing it would not provide any novel contribution to our understanding), but that signallers, even honest ones, have to pay an extra cost for honesty (i.e. a strategic cost), that is, signals have to be 'wasteful' at the equilibrium (hence the name of the theory, i.e. handicap principle, see the previous Grafen (1990) quote; and hence the debate that followed the original introduction of the idea; see Maynard Smith 1976). Zahavi & Zahavi (1997) explicitly claimed that exactly because honest signalling selects for waste, this kind of selection is different from traditional Darwinian natural selection and, accordingly, called it 'signal selection'.

This Zahavi–Grafen conclusion was soon reinforced by Maynard Smith (1991) investigating a discrete action–response model, the so-called 'Sir Philip Sidney game'. The results of this model confirmed Grafen's (1990) conclusion since Maynard Smith (1991) was able to show that as long there is a conflict of interest between the signaller and the receiver, signals have to be costly to be honest (i.e. there will be a strategic cost paid at the equilibrium by honest signallers). It follows that in cases of shared interest, signals can be cost-free (Maynard Smith 1991), which became a widely known and widely accepted conclusion.

The Zahavi–Grafen conclusion, that honest signallers have to pay an extra cost (handicap) at the equilibrium for the maintenance of honesty, is not correct simply because it is not this cost that maintains honesty (Hurd 1995; Számadó 1999; Lachmann et al. 2001). Potential cheaters will refrain from cheating only if the potential cost of cheating is higher than the potential benefits resulting from the action. Honest signallers will pay an extra cost at the equilibrium only if a biological constraint exists that creates a link between the costs paid by honest and dishonest signallers (Számadó 1999; Lachmann et al. 2001). In other words, the existence of strategic cost (handicap) at the equilibrium is not a theoretical necessity, as claimed by Zahavi and Grafen, but the result of biological constraints. In the absence of such constraints, signals need not be costly to be honest not even under conflict of interest (i.e. there is no need for strategic cost paid by honest signallers even if interests are in conflict). The existence of such a constraint is an empirical question and without this knowledge the equilibrium signal cost paid by honest signallers cannot be predicted. In the light of the importance of such a constraint, it is easy to interpret the theoretical results.

Grafen (1990) and Maynard Smith (1991) arrived at the conclusion that honest signalling necessarily involves waste simply because they implicitly assumed a constraint between the two types of cost. Hurd (1995) was the first to investigate a model in which no such constraint was assumed. Not surprisingly, he was able to show that the cost of signals at the equilibrium for honest signallers (i.e. the realized cost) can be zero or even negative if the cost of cheating is high enough. Next, Számadó (1999) was able to show in a general version of these action–response games that the source of discrepancy between the results of Maynard Smith (1991) and Hurd (1995) is the assumption on signal cost. If one assumes a constraint (such as Maynard Smith did) then signals under conflict of interest have to be costly; if not (such as Hurd did) then they need not be (see Appendix 2 and Fig. A1 for a detailed explanation).

A few years later, Lachmann et al. (2001) were able to show that the same logic applies to continuous action–response games as well. If the cost of signals can vary with the quality being signalled, then a cost function can be found that gives an arbitrarily close to zero cost at the equilibrium; if not, then there will necessarily be a positive cost at the equilibrium. Lachmann et al. (2001) constructed such 'cost-free' cost functions for the models of Grafen (1990) and Godfray (1991) showing that the result of Hurd (1995) is not an artefact of discreteness; that relaxing the implicit constraint in Grafen's (1990) model between the two types of cost will release honest signallers from the necessity to pay a handicap at the equilibrium; and thus discrete and continuous games follow the same logic.

In terms of symmetric games it was Enquist (1985) who was able to show that honest signalling of intentions can be evolutionarily stable, provided that the cost of fighting for weak individuals against strong ones exceeds a given threshold. While signals in Enquist's game are cost-free by definition, many authors, including Grafen (1990), took the result as support for the handicap principle. This somewhat conflicting interpretation comes from the fact that while signals are cost-free, signallers can get into a fight, and fighting is costly and more so for weak individuals. Accordingly, the results of Enquist's model were often interpreted as support for the claim that threat or dominance displays have to be 'vulnerability handicaps' or 'risk handicaps' (Grafen 1990; Vehrencamp 2000). However, the basic principle described in the case of action–response games works here too. It is not the realized cost of the honest signallers that maintains honesty but the potential cost of cheating. In terms of Enquist's model, it is the cost for a weak individual fighting a strong one that has to be higher than the potential gains, and the cost for a strong individual fighting another strong one can take any value; it can be zero or even negative. Thus, there is no need for any additional cost to maintain honesty at the equilibrium; in fact if a costly signal (handicap) is introduced in an obligatory way then it is the weak individuals who will suffer this extra cost and not the strong ones (Hurd 1997). If, however, cost-free or negative-cost signals are allowed to invade, they will spread and replace handicaps (Számadó 2003). This follows exactly because it is not the realized cost of honest individuals that maintains honesty, and if these signals are allowed to evolve then the most cost-efficient ones (the ones with zero or even negative costs) will replace the less cost-efficient handicaps. All in all, threat or dominance displays can be truly and fully cost-free (or even with negative costs) for honest individuals as long as the potential cost of cheating is high enough (Számadó 2003, 2008). This potential cost may come from the optimization cost inherent in the situation (i.e. for a weak individual it is not worth fighting a strong one) and thus it need not impose a realized signal cost for honest signallers at the equilibrium.

All in all, both action–response and symmetric games show that it is not the cost paid by the honest signallers that maintains the equilibrium but the potential cost of cheating. Whether this potential cost implies a realized cost for honest signallers depends on the structure of the game; in general it depends on the biological constraints of the system. Strategic signal cost paid at the equilibrium (i.e. handicap) by honest signallers is not a theoretical necessity as predicted by Zahavi (1975, 2008) and Grafen (1990), and it cannot be predicted without the knowledge of these constraints. In other words, adding a strategic cost to equilibrium signals (i.e. handicapping them) is not the only way to create a high potential cost of cheating. This potential cost can come from several sources: (1) it can be simply the efficacy cost of a signal (e.g. not everyone has the resources to build a lighthouse); (2) it can be a strategic cost imposed on potential cheaters by honest signallers; or (3) it can be a cost paid because the out-of-equilibrium signal intensity indicates a suboptimal decision in the given context (i.e. optimization cost). All three types of cost can maintain honest signalling; the difference lies in the realized cost paid by honest signallers at the equilibrium. Optimization and efficacy costs will result in cost-free or minimal-cost signals, while strategic cost, depending on the presence of biological constraints, can result in all kinds of signals (i.e. cost-free, minimal-cost and handicap). Accordingly, all types of signal are possible at the equilibrium even if there is a conflict of interest between signaller and receiver. Table 1 summarizes all the options.

IMPLICATIONS FOR EXPERIMENTAL TESTING

What are the implications for empirical researchers? Researchers, instead of trying to measure equilibrium signal cost, should ask what the potential costs of cheating in a given situation are, and whether these potential costs should necessarily impose an equilibrium cost on the honest signallers. Below, I discuss many examples where the potential cost need not impose equilibrium cost on honest signallers. Here, I want to highlight the problems related to measuring equilibrium (realized) cost. First, measuring equilibrium (realized) cost is not informative simply because it is not the realized cost at the equilibrium that maintains the honesty of signals (Lachmann et al. 2001). The prediction of honest signalling theory is that if it is the signal cost that maintains honesty, then the marginal cost and marginal benefits coming from signalling have to be equal at equilibrium (Lachmann et al. 2001; Bergstrom et al. 2002). However, these marginal costs and benefits can be measured only by manipulating the animals into giving out-of-equilibrium signals and manipulating the receivers to react to these signals (Kotiaho 2001; Roberts et al. 2004; Moreno-Rueda 2007). The second reason why measuring the realized (equilibrium) cost is not informative is that, despite the recognized importance of distinguishing efficacy cost from strategic cost (Grafen 1990; Maynard Smith & Harper 1995, 2003), no one has yet described how to separate the two kinds of cost in experiments. If honest signallers pay for handicaps then the realized cost at the equilibrium is the sum of the efficacy and strategic (handicap) costs. Obviously, to be able to measure the handicap component, researchers would have to separate the two types of cost. Yet, as said, no such method has been suggested. Researchers have mostly compared resting metabolic rates (RMR) to active metabolic rates (AMR) or to daily energy budgets (McCarty 1996; Oberweger & Goller 2001). Note that this will only measure the realized cost and it will not tell us what proportion, if any, of this cost is the ‘handicap’, and what proportion is the efficacy cost. Of course, if there is no increase in AMR compared to RMR, then there cannot be a strategic cost involved; however, even if there is an increase in cost, this may be only the efficacy cost (i.e. the minimum necessary investment to transmit the given information reliably). As long as a clear methodology is not available to distinguish efficacy cost from handicaps, then even measuring a positive realized cost at the equilibrium cannot be taken as confirmation of the predictions of the handicap principle.

In summary, measuring the realized cost is not informative for two reasons: (1) on the one hand, measuring zero realized cost is no proof that signal cost plays no role in maintaining honesty since it is not the realized cost that maintains it; (2) on the other hand,

Table 1
Signal types at the equilibrium as a function of the potential cost of cheating, with mechanisms and examples

Type	Potential cost of cheating			Constraint on signal cost function?	Realized cost of honest signallers	Observed signal at the equilibrium	Examples of mechanisms	Examples of context
	Efficacy cost	Optimization cost	Strategic cost					
a	Negligible	Yes	No	No	Negligible	Cost-free signal	Shared interest	Mate recognition, alarm calls, group cohesion
b	Yes	Yes	No	No	Efficacy cost	Minimal-cost signal	Pooling equilibria	Begging?
							Indices	Territorial displays, dominance displays, sexual displays
							Amplifiers	Sexual displays, dominance displays
							Individual recognition	Dominance displays
							Frequency-dependent selection	Sexual displays, aposematic displays
							Proximity risk	Threat displays
							Inclusive fitness	Begging
							Punishment of cheaters	Food discovery calls, aposematic displays
c	Yes	No	No	No	Efficacy cost	Minimal-cost signal	Performance displays	Assessment behaviour during contest
d	Negligible	No	Yes	No	Negligible or strategic cost	Cost-free signal or handicap		
e	Yes	No	Yes	No	Efficacy cost or efficacy cost and strategic cost	Minimal-cost signal or handicap		
f	Negligible	No	Yes	Yes	Strategic cost	Handicap		
g	Yes	No	Yes	Yes	Efficacy cost and strategic cost	Handicap		

Examples are listed only for the cases discussed in the article, i.e. where high potential cost of cheating can be maintained without imposing realized cost on honest signallers at the equilibrium. Optimization, efficacy and strategic costs can all maintain honest signalling with the difference lying in the realized cost paid by honest signallers at the equilibrium. Optimization and efficacy costs will result in cost-free or minimal-cost signals (types a, b and c, respectively), while strategic cost, depending on the presence of constraints (on the mechanism level), can result in all kinds of signals: in the lack of constraint, the result is either a cost-free/minimal-cost signal or handicap, depending on the magnitude of the efficacy cost, see types d and e; on the other hand, if there is a constraint, then the result is a handicap: types f and g.

measuring positive realized cost is no proof that the given trait is a handicap since the realized cost is the sum of efficacy and handicap costs and, at the moment, there is no methodology to separate the two.

The problem is further confounded with the fact that energy expenditure need not equal cost in evolutionary terms (Kotiaho 2001; Moreno-Rueda 2007). Thus, for example, Moreno-Rueda (2007) argued that measuring the effect of begging on growth rate is more informative then quantifying energy expenditure, yet only a handful of studies have followed this protocol. All reviews (Kotiaho 2001; Roberts et al. 2004; Moreno-Rueda 2007) agree that to reach a conclusion, experiments with better designs are required: either the signal trait should be manipulated to known condition/quality levels or signal trait and condition/quality should be manipulated simultaneously, and in both cases both the marginal benefit and marginal cost resulting from the manipulation of the signal trait should be measured. Conclusions of the theoretical models suggest, however, that a more promising approach is to search for potential constraints that can link the costs paid by potential cheaters and honest signallers. If one can show that such a mechanism exists in a given situation, then honest signals have to have high realized cost at the equilibrium (i.e. they will be handicaps); if not, then honest signals can be low cost or even cost-free.

MECHANISMS OF HONEST SIGNALLING THAT PROMOTE LOW REALIZED COST FOR HONEST SIGNALLERS

My aim here is not to provide an exhaustive list, but just to show examples of how potential cost need not impose an extra realized cost (handicap) on honest signallers (hence those mechanisms of honest signalling that potentially impose realized cost on honest signallers will not be discussed here).

(1) Shared interest in itself can guarantee the honesty of a signalling system (Maynard Smith 1991). There is no need for any signal cost simply because the optima of the signaller and the receiver are already in accord. Mate recognition calls, alarm calls or calls maintaining group cohesion can fall into this category. Perhaps here it is most obvious that optimization cost can maintain honest signalling. The potential cost of cheating comes from the cost of making a suboptimal decision. While not always apparent, the principle of optimization cost is the same; as we will see, such costs can exist (depending on the biological details of the system) even if there is a conflict of interest between signaller and receiver.

(2) Pooling equilibria. Shared interest need not be complete to allow cost-free honest signalling. If there is a distribution of signaller types and the two ends of the distribution would prefer to inform the receivers of their states rather than hiding them, then so-called ‘pooling’ equilibria can evolve (Lachmann & Bergstrom 1998) in which signallers in different states will give the same signal, but, importantly, these groups will be continuous with regard to quality. In other words, these equilibria are only partially informative but honest in the sense that the best and the worst signaller types will use different signals. Importantly, signals at these pooling equilibria can be cost-free. The potential cost of cheating comes from the fact that receivers take the quality of the signaller as the average of the pooling groups and at the equilibrium it is not worth any signaller pretending to be the member of any other group (because the change in perceived value is discrete, corresponding to the group averages, and does not change continuously with the signal). Begging has been suggested as a potential context for such signals (Bergstrom & Lachmann 1998; but see Brilot & Johnstone 2003). Unfortunately, the idea has never been taken up and no experimental data are available.

(3) Indices are widespread both in animal and in human communication. An index is a signal whose reliability is maintained by a mechanistic link (physical connection) between signal intensity and a given trait (Maynard Smith & Harper 1995). The source of this mechanistic link can be very diverse ranging from the presence of a territory owner and its ability to countermark (i.e. scent countermarking, Rich & Hurst 1999), through physical correlation between the length of the vocal tract and its formants (Reby & McComb 2003; Sanvito et al. 2007), to the pleiotropic effects of genes such as the melanocortin system in vertebrates in which melanocortins affect coloration, aggressiveness, sexual activity and immune response (Ducrest et al. 2008). Indices can be found in all possible contexts ranging from sexual selection to aggressive communication.

(4) Performance displays can be thought of as ‘behavioural indices’ (although Hurd & Enquist 2005 classified both morphological and behavioural indices as subsets of ‘performance signalling’). Here, reliability is maintained by a physical correlation between condition and performance. Of course, performance can be costly but it is not the cost that matters. The animal carrying out the same action more cheaply is not cheating; it is just more efficient. Examples include tail beating in cichlid fish, *Nannacara anomala*, during contests (Enquist et al. 1990) or the opercular display of Siamese fighting fish, *Betta splendens*, in which the fish practically ‘holds its breath’ to test the condition of its opponent (Abrahams et al. 2005).

(5) Amplifiers do not create a correlation but they help the receiver to assess the quality of the signaller; that is, amplifiers increase the information content of the original cue or signal (Hasson 1997). Amplifiers are expected to evolve when it is in the interest of high-quality signallers to differentiate themselves from low-quality ones (Taylor et al. 2000). They can be of low cost or cost-free (Taylor et al. 2000). Examples include the dark central patch on the abdomen of the jumping spider, *Plexippus paykulli* (Taylor et al. 2000) and the gaping displays of collared lizards (*Crotaphytus* spp., Lappin et al. 2006).

(6) Individual recognition. The ability to recognize other members of the species individually allows the evolution of cheap signals in many contexts (Tibbetts & Dale 2007). Here, honesty is maintained by the unity that we call ‘individual’. Hence information gathered by some means, such as play fighting, eavesdropping or assessment, earlier on could be reused later without paying the same cost. Of course, it is most useful in those cases where the given quality is more or less stable during reasonably long enough periods relevant to the individuals. However, there are good numbers of such traits, including resource-holding potential (Parker 1974), position in a group hierarchy, propensity to cooperate or genetic quality. Signals that make individual recognition possible are expected to be cheap (Dale et al. 2001). Examples include grunts and ‘girneys’ of female rhesus monkeys, *Macaca mulatta*, as signals of benign intent (Silk et al. 2000), maintenance of a dominance hierarchy in paper wasps, *Polistes dominulus* (Tibbetts & Dale 2004) and alarm calls of marmots, *Marmota flaviventris*, in which individuals judge the reliability of the call based on the identity of the signaller (Blumstein et al. 2004).

(7) In the case of frequency-dependent selection, reliability is maintained by a negative feedback between the frequencies of the different types of strategy and their payoffs. One can argue that this type of mechanism cannot maintain a fully honest system as there will always be cheaters present. The frequency of cheaters, however, can be arbitrarily close to zero in such systems, as it all depends on the payoffs of the different types of strategy. Examples include the different life strategies of male bluegill sunfish, *Lepomis macrochirus* (Dominey 1980; Gross & Charnov 1980), cleaner fish

feeding on host mucus (Freckleton & Cote 2003) and Batesian mimicry (Joron 2003).

(8) Proximity risk. In the case of threat displays, it is proximity to an opponent ready to retaliate that creates reliability. Accordingly, threat displays are expected to be honest only within a certain distance from the opponent, the so-called 'honest striking distance' which is a function of resource value, weaponry and species-specific fighting technique (Számadó 2008). This commitment creates only a potential cost (i.e. risk) and is not an equilibrium cost paid by everyone (i.e. not a handicap). There is no need for additional 'vulnerability handicaps' as proposed by Zahavi & Zahavi (1997). In fact, the form of all known threat displays prepares the animal to fight (i.e. decreasing this potential risk further), as it consists of a presentation of weapons in ready-to-fight form (Walther 1984; Számadó 2003).

(9) Punishment of the cheaters (social cost). Here, honesty is maintained by the punishment of the cheaters. Thus, the marginal cost of deviating from the honest equilibrium comes from the potential risk of being detected and punished. Again, this is a potential cost and need not be present at the equilibrium (Lachmann et al. 2001). Examples include increased harassment of monkeys who do not give a food call upon discovering a food item (Hauser 1992) and selective removal of nontoxic individuals by predators from the population in the case of automimic insects (Gamberale-Stille & Guilford 2004).

(10) Loss of inclusive fitness can be another potential cost. Johnstone (1999) was able to show in a competitive model of begging that high signaller-to-signaller relatedness, assuming more than two signallers, can result in low-cost equilibrium signals. Here again, an honest signaller optimizing its inclusive fitness need not suffer an extra signal cost at the equilibrium. A review of the experimental literature also supports the importance of inclusive fitness costs in begging (Moreno-Rueda 2007).

All in all, signals will be honest under conflict of interest if there is a mechanism that can create a correlation between the signal and the advertised trait and that can guarantee that the marginal cost of producing equilibrium signals is higher than the marginal benefit gained from them. As we have seen, there is a range of mechanisms that can create these conditions without imposing an extra signal cost (handicap) at the equilibrium on honest signallers. Table 1 summarizes the possible cases.

CONCLUSIONS

The honesty of communication is maintained by the potential cost of cheating and not by the cost paid by honest signallers at the equilibrium. Whether this potential cost imposes a cost on honest signallers depends on the biological details of the system and its presence and magnitude cannot be predicted by theoretical models a priori without knowing the relevant details. There are several mechanisms (see the examples reviewed above) that can achieve such high potential cost without imposing realized equilibrium cost on honest signallers. It is not only handicapping equilibrium signals that ensure a high potential cost of cheating. Accordingly, strategic cost (i.e. handicap), contrary to the Zahavi–Grafen claim, is neither a necessary nor a sufficient condition for honest signalling and thus honest signals need not be costly, even under conflict of interest. In other words, the central tenet of the Zahavi–Grafen handicap principle, that strategic cost paid at the equilibrium by honest signallers is a necessary condition for honest signalling, is a fallacy. Note that handicaps may still exist in nature and may be plentiful; it just means that the presence of handicaps at the equilibrium is not the theoretical and logical necessity, even under conflict of interest, that was predicted by the handicap principle. It follows that it is also a fallacy to claim that all signals (Zahavi 2007, 2008; Zahavi & Zahavi 1997) or even a subset of signals (i.e. signals used in the case of conflict of interest, Grafen 1990) have to be handicaps.

The operational notion of strategic cost is questionable itself as long as there is no solid methodology to distinguish efficacy cost from strategic cost. Measuring equilibrium cost is not informative for two reasons: first, it is not the equilibrium cost that maintains honesty, and, second, the above-mentioned lack of empirical methodology to separate strategic cost from efficacy cost. Only measuring the marginal costs and marginal benefits of signals can tell whether honesty is maintained by a signal cost or not. Research effort could be better spent on identifying the proximate mechanisms of honest signalling and thus identifying those potential sets of constraints that might create a link between the potential cost of cheaters and the realized cost of honest signallers, which in turn might lead to costly equilibria.

There is already a long list of mechanisms (reviewed in the previous section) that lack such constraints; thus, these mechanisms can maintain honesty (i.e. high potential cost for cheaters) without imposing an extra signal cost at the equilibrium on honest signallers. The existence of other such mechanisms cannot be ruled out a priori. The diversity of honest signalling and of the mechanisms that maintain it, producing cost-free, minimal-cost signals and handicaps at the equilibrium, will reveal itself in its full shape and beauty when researchers are primed to look for it.

Acknowledgments

I thank László Garamszegi and two anonymous referees for their helpful comments. This work was supported by the Hungarian National Science Foundation (OTKA, Grant T049692).

References

Abrahams, M. V., Robb, T. L. & Hare, J. F. 2005. Effect of hypoxia on opercular displays: evidence for an honest signal? *Animal Behaviour*, **70**, 427–432.

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

Bergstrom, C. T., Számadó, Sz. & Lachmann, M. 2002. Separating equilibria in continuous signalling games. *Philosophical Transactions of the Royal Society B*, **357**, 1595–1606.

Blumstein, D., Verneyre, L. & Danial, J. C. 2004. Reliability and the adaptive utility of discrimination among alarm callers. *Proceedings of the Royal Society B*, **271**, 1851–1857.

Brilot, B. O. & Johnstone, R. A. 2003. The limits to cost-free signalling of need between relatives. *Proceedings of the Royal Society B*, **270**, 1055–1060.

Dawkins, M. S. & Guilford, T. 1991. The corruption of honest signalling. *Animal Behaviour*, **41**, 865–873.

Dale, J., Lank, D. B. & Reeve, H. K. 2001. Signaling individual identity versus quality: a model and case studies with ruffs, queleas, and house finches. *American Naturalist*, **158**, 75–86.

Dominey, W. J. 1980. Female mimicry in male bluegill sunfish: a genetic polymorphism? *Nature*, **284**, 546–548.

Ducrest, A., Keller, L. & Roulin, A. 2008. Pleiotropy in the melanocortin system, coloration and behavioural syndromes. *Trends in Ecology & Evolution*, **23**, 502–510.

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

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

Freckleton, R. P. & Cote, I. M. 2003. Honesty and cheating in cleaning symbioses: evolutionarily stable strategies defined by variable pay-offs. *Proceedings of the Royal Society B*, **270**, 299–305.

Gamberale-Stille, G. & Guilford, T. 2004. Automimicry destabilizes aposematism: predator sample-and-reject behaviour may provide a solution. *Proceedings of the Royal Society B*, **271**, 2621–2625.

Getty, T. 1998. Handicap signalling: when fecundity and viability do not add up. *Animal Behaviour*, **56**, 127–130.

Getty, T. 2006. Sexually selected signals are not similar to sports handicaps. *Trends in Ecology & Evolution*, **21**, 83–88.

Godfray, H. J. C. 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**, 246–517.

Gross, M. R. & Charnov, E. L. 1980. Alternative male life histories in bluegill sunfish. *Proceedings of the National Academy of Sciences, U.S.A.*, **77**, 6937–6940.

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

Hauser, M. D. 1992. Cost of deception: cheaters are punished in rhesus monkeys (*Macaca mulatta*). *Proceedings of the National Academy of Sciences, U.S.A.*, **89**, 12137–12139.

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

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

Hurd, P. L. & Enquist, M. 2005. A strategic taxonomy of biological communication. *Animal Behaviour*, **70**, 1155–1170.

Johnstone, R. A. 1999. Signaling of need, sibling competition, and the cost of honesty. *Proceedings of the National Academy of Sciences, U.S.A.*, **96**, 12644–12649.

Joron, M. 2003. Mimicry. In: *Encyclopedia of Insects* (Ed. by R. T. Cardé & V. H. Resh), pp. 714–726. New York: Academic Press.

Kotiaho, J. 2001. Cost of sexual traits: a mismatch between theoretical considerations and empirical evidence. *Biological Review*, **76**, 365–376.

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

Lachmann, M., Számadó, Sz. & Bergstrom, C. T. 2001. Cost and conflict in animal signals and human language. *Proceedings of the National Academy of Sciences, U.S.A.*, **98**, 13189–13194.

Lappin, A. K., Brandt, Y., Husak, J. F., Macedonia, J. M. & Kemp, D. J. 2006. Gaping displays reveal and amplify a mechanically based index of weapon performance. *American Naturalist*, **168**, 100–113.

McCarty, J. P. 1996. The energetic cost of begging in nestling passerines. *Auk*, **113**, 178–188.

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

Maynard Smith, J. 1976. Sexual selection and the handicap principle. *Journal of Theoretical Biology*, **57**, 239–242.

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

Maynard Smith, J. & Harper, D. G. C. 2003. *Animal Signals*. Oxford: Oxford University Press.

Moreno-Rueda, G. 2007. Is there empirical evidence for the cost of begging? *Journal of Ethology*, **25**, 215–222.

Oberweger, K. & Goller, F. 2001. The metabolic cost of birdsong production. *Journal of Experimental Biology*, **204**, 3379–3388.

Parker, G. A. 1974. Assessment strategy and the evolution of animal conflicts. *Journal of Theoretical Biology*, **47**, 223–243.

Rich, T. & Hurst, J. L. 1999. The competing countermarks hypothesis: reliable assessment of competitive ability by potential mates. *Animal Behaviour*, **58**, 1027–1037.

Reby, D. & McComb, K. 2003. Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour*, **63**, 519–530.

Roberts, M. L., Buchanan, K. L. & Evans, M. R. 2004. Testing the immunocompetence handicap hypothesis: a review of the evidence. *Animal Behaviour*, **68**, 227–239.

Sanvito, S., Galimberti, F. & Miller, E. H. 2007. Vocal signalling of male southern elephant seals is honest but imprecise. *Animal Behaviour*, **73**, 287–299.

Searcy, W. A. & Nowicki, S. 2005. *The Evolution of Animal Communication*. Princeton, New Jersey: Princeton University Press.

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

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

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

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

Számadó, Sz. 2008. How threat displays work: species-specific fighting techniques, weaponry and proximity risk. *Animal Behaviour*, **76**, 1455–1463.

Taylor, P. W., Hasson, O. & Clark, D. L. 2000. Body postures and patterns as amplifiers of physical condition. *Proceedings of the Royal Society B*, **267**, 917–922.

Tibbetts, E. A. & Dale, J. 2004. A socially-enforced signal of quality in a paper wasp. *Nature*, **432**, 218–222.

Tibbetts, E. A. & Dale, J. 2007. Individual recognition: it is good to be different. *Trends in Ecology & Evolution*, **22**, 529–537.

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

Walther, F. R. 1984. *Communication and Expression in Hoofed Mammals*. Bloomington, Indiana: Indiana University Press.

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

Zahavi, A. 1977. The cost of honesty. *Journal of Theoretical Biology*, **67**, 603–605.

Zahavi, A. 2007. Sexual selection, signal selection and the handicap principle. In: *Reproductive Biology and Phylogeny of Birds* (Ed. by B. G. M. Jamieson), pp. 143–159. Enfield, New Hampshire: Science Publishers.

Zahavi, A. 2008. The handicap principle and signalling in collaborative systems. In: *Sociobiology of Communication* (Ed. by P. d'Ettore & D. P. Hughes), pp. 1–11. Oxford: Oxford University Press.

Zahavi, A. & Zahavi, A. 1997. *The Handicap Principle*. New York: Oxford University Press.

APPENDIX 1: GLOSSARY

Cue: a feature of the external world that can decrease the uncertainty in the underlying decision problem. Cues can be nonliving features or part of a living organism but importantly they did not evolve to fit this function (Maynard Smith & Harper 1995).

Dishonest signal: a signal that hides the relevant quality of the signaller from the receiver; the receiver consequently treats the signaller as if it belongs to a different class. Need not imply intentionality.

Efficacy cost: necessary signal cost to transmit a given piece of information (Maynard Smith & Harper 1995, 2003).

Equilibrium cost: the cost of equilibrium signals, i.e. signal cost paid at the equilibrium (includes both efficacy and strategic costs).

Evolutionary stability: long-term stability, technically given as the evolutionarily stable strategy.

Evolutionarily stable strategy (ESS): a strategy (a morphological feature or behaviour) that, if played by all members of the population, prevents a mutant strategy invading the population (Maynard Smith 1982); in other words it is not worth any player deviating from it.

Handicap principle: a theory that predicts that the honesty of signals is maintained by the strategic cost (handicap) paid at the equilibrium (Zahavi 1975; Grafen 1990).

Honest signal: a signal that reveals the relevant quality of the signaller to the receiver. Need not imply intentionality.

Marginal benefit: change in benefits resulting from the use of the signal as a function of the change in signal intensity; more formally, the marginal benefit is the derivative of the benefit function as a function of signal intensity.

Marginal cost: change in signal cost as a function of the change in signal intensity; more formally, the marginal cost is the derivative of the signal cost function as a function of signal intensity.

Pooling equilibrium: partially informative signalling equilibrium; continuous pools of signallers give a signal of the same intensity.

Separating equilibrium: fully informative signalling equilibrium; signallers with different levels of quality give signals of different intensity.

Signal: a feature that can potentially decrease the uncertainty in the underlying decision problem, and that evolved because, on average, it increases the fitness of the signaller by altering the behaviour of the receiver.

Signal cost: cost of signal as a function of quality; it may or may not be the same for all levels of quality.

Signalling game: a game in which a player with private information (the signaller) may or may not send a signal to the receiver to reveal the private information (potentially useful to the receiver) before the receiver makes a decision that could impact the fitness of both players.

Strategic cost: signal cost on top of the efficacy cost and paid at the equilibrium; suggested by the handicap principle as a necessary and sufficient condition for honest signalling (Zahavi 1975; Grafen 1990).

APPENDIX 2: COST-FREE SIGNALLING EQUILIBRIUM IN DISCRETE ACTION–RESPONSE GAMES

Hurd (1995) was the first to show that honest cost-free signalling is possible in discrete action–response games even under conflict of interest. The most general case is when the signaller's fitness depends on the signaller's state and the fitness of both players is influenced by the survival of the other. If we assume a discrete model with two states, two signals and two responses, the conditions of evolutionarily stable honest signalling are as follows (Számadó 1999):

$$W_h + rV_h > 0 \quad (1)$$

$$W_l + rV_l < 0 \quad (2)$$

$$V_h + rW_h > C_h \quad (3)$$

$$V_l + rW_l < C_l \quad (4)$$

$$V_h + rW_h > 0 \quad (5)$$

$$V_l + rW_l > 0 \quad (6)$$

where W, V and C denote the receiver's fitness, the signaller's fitness and the cost of signalling, respectively, and l and h denote the quality of the signaller ('high' and 'low', respectively). The fitness of each player can be influenced by the survival of the other player (r). The first two inequalities describe the receiver's conditions for honest signalling, the second two are the signaller's conditions, and the last two specify the conflict of interest. Reversing the last inequality would mean that there is no conflict of interest between signaller and receiver. It is clear that if $r = 1$, assuming that signalling is beneficial for the receiver, then there can be no conflict of interest (since $V_l + rW_l = W_l + rV_l$; thus inequality 6 cannot be fulfilled). This implies that in this case C_l can be zero or less than zero (since the left-hand side of inequality 4 need not be greater than zero). However, if $0 < r < 1$ then at least C_l should be greater than zero (inequality 4). That is, signalling for low-quality individuals must be costly when there is a conflict of interest. Signalling for high-quality individuals, however, need not be costly even in this case (inequality 3). Since at the honest equilibrium only high-quality individuals signal, the observed equilibrium cost can be zero or even negative. Figure A1a depicts the regions of honest signalling when there is a conflict of interest. Green denotes the region where honest signalling is an ESS (note that C_h can be zero or even less than zero), whereas the region with stripes denotes the area specified by Grafen's (1990) main handicap results (i.e. that signals are costly $C_l > 0$ and $C_h > 0$, and costlier for worse signallers $C_l > C_h$). We can see that while there is an overlap between the two regions, the overlap is far from complete. Note that none of the regions is a subset of the other; hence, we can see that Grafen's conditions are neither necessary nor sufficient conditions for honest signalling.

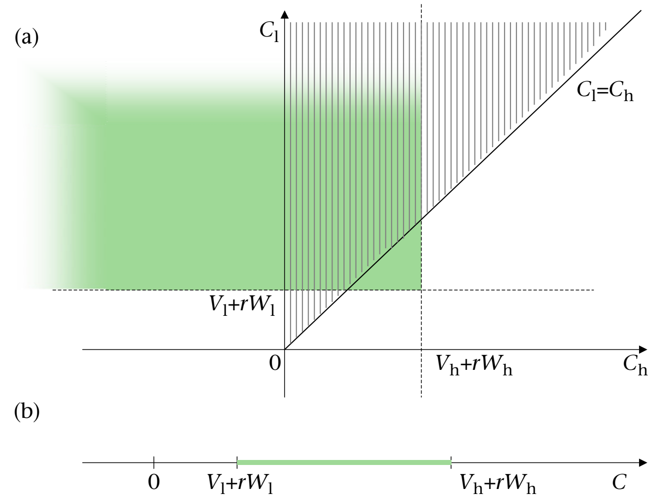

Figure A1. (a) Regions of honest evolutionarily stable signalling (green, note that this region is open-ended towards high C_l and low C_h values): when $C_h \neq C_l$, $V_h \neq V_l$ and $r > 0$. Note C_l must be greater than zero, whereas C_h can be zero or even smaller than zero. (b) Regions of honest evolutionarily stable signalling when $C_h = C_l$ (thick green line); note both C_h and C_l must be greater than zero (after Számadó 1999).

If, on the other hand, we assume that $C_l = C_h$ then both costs should be greater than zero for honest signalling to be stable. Figure A1b depicts this case. This corresponds to the Sir Philip Sidney game investigated by Maynard Smith (1991). Note that Grafen's main handicap result describes the condition of honest signalling only in this case. Which case (i.e. Fig. A1a or b) holds for a given biological system is determined by the details of that system, that is, by the biological constraints. Accordingly, strategic cost paid at the equilibrium is not a theoretical necessity as predicted by the handicap principle (even under conflict of interest).

The same general logic applies to continuous models as well (Lachmann et al. 2001); thus, when the signal cost function is a function of the quality being signalled, it is possible to find 'cost-free' equilibria where the equilibrium cost of signals for honest signallers is negligible (i.e. arbitrarily close to zero, see Lachmann et al. 2001); if, on the other hand, the signal cost function does not depend on the quality being signalled, then the equilibrium cost for honest signallers will necessarily be greater than zero.