J. theor. Biol. (1999) **198,** 593–602
Article No. jtbi.1999.0935, available online at http://www.idealibrary.com on IDEAL

The Validity of the Handicap Principle in Discrete Action–Response Games

Szabolcs Számadó*

Department of Plant Taxonomy and Ecology, Eötvös Lorand University, Ludovika tér 2, *H*-1083 *Budapest, Hungary*

(*Received on* 18 *August* 1998, *Accepted in revised form on* 16 *March* 1999)

The validity of the handicap principle has spawned much debate in spite of the existence of a formal treatment. Simple models constructed to further investigate the issue were able both to prove and to disprove some of its claims. Here I show with the aid of a more general model, which takes into account both assumptions presented in these previous simple models: (1) that the previous results are not in conflict since they can be obtained as specific cases of this general model; (2) that ESS communication need not use costly signals, that is, even under conflict of interest, the cost of a signal used by a high-quality individual can be zero (or even negative) provided that the cost for low-quality signallers is high enough; (3) that only the cost relative to the benefits of the interaction should be higher for worse signallers; and (4) that in a discrete model the differential cost is only a necessary but not a sufficient condition for evolutionarily stable reliable communication.

1. Introduction

There is a long raging debate about the nature of communication in biological systems with two conflicting traditions. According to the first school, cheating is widespread among animals and they communicate in order to manipulate each other (Krebs & Dawkins, 1984). On the contrary, the second school puts the stress upon the honesty of the communication which is maintained by the cost of signals: this is the so-called "handicap principle" (Zahavi, 1975, 1977). Although this later view has gained theoretical support (Nur & Hassen, 1985; Pomiankowski, 1987; Grafen, 1990) it is still widely debated either on the ground of flawed assumptions: for instance, the receiver's cost is neglected (Dawkins & Guilford, 1991), or more generally, because it is not clear how specific is the case represented by Grafen's model. Although serious efforts were made to clarify the situation with simple models, the conflict remains. Shortly after the publishing of Grafen's result Maynard Smith (1991) constructed the Philip Sidney game to make it easier to follow Grafen's argument. He was able to show that in the case of conflict of interest, honest signals must be costly, but cost-free signals can be honest if there is no such conflict between signaller and receiver. On the contrary, in a more recent model, Hurd (1995) has shown that even under conflict of interest, cost-free communication can be honest provided that the cost, paid by the mutant who does not play the ESS, is large enough. Since both models consider the simplest case of signalling (only two states, two actions and two responses are allowed) it would be interesting to know the underlying differences in the

* Author to whom correspondence should be addressed. E-mail: szamszab@ludens.elte.hu.

0022–5193/99/013593 + 10 $30.00/0

assumptions resulting in such conflicting conclusions. In order to investigate the issue, here I propose a more general model by which the previous ones can be explained. It considers both assumptions presented in the previous models; that is, the possible dependence of the signaller's fitness on the signaller state and the possible influence of the fitness of each participant on the survival of the other.

In the following section the general model is introduced. In Section 3 three conditions necessary for evolutionarily stable communication under conflict of interest are described and the cost of signalling is defined. The analyses follows in Section 4. Finally, Section 5 concludes.

2. The Action–Response Game

The model concerns a minimal signalling game with discrete signals. For the sake of simplicity and continuity, it is couched in terms of the Basic Action–Response game introduced by Hurd (1995). I will use his notation wherever possible. There are two players: a signaller and a receiver. The signaller can be either in a High or in a Low state, $z := \{H, L\}$, which is decided by nature. This initial move is hidden from the receiver; it only observes the behaviour the signaller has chosen, Signal or Not to signal, $p := \{S, N\}$, and then, based on this observation makes its reply: accepts—turns the signaller Up, or rejects—turns it Down, $q := \{U, D\}$. In the most general case the following assumptions can be made about the signaller's and the receiver's payoffs:

(1) The receiver's fitness (L_r) depends both on the state of the signaller (z) and on the receiver's response (q):

$$L_r = W(z, q). \quad (1)$$

(2) The signaller's fitness (L_s) is a sum of the value of the receiver's response (V) and the cost of signalling (C):

$$L_s = V(z, q) - C(z, p) \quad (2)$$

where V depends on both the state of the signaller (z) and on the receiver's reply (q), whereas C depends on the state of the signaller (z) and on the behaviour of the signaller (p).

(3) The fitness of each player can be influenced by the survival of the other player (r) which can mean, for instance, that they are relatives, or in the same way can help each other (Maynard Smith, 1991). Based on these assumptions the inclusive fitness of the signaller (E_s) and the receiver (E_r) can be written as follows:

$$E_s(z, p, q) = L_s + rL_r$$
$$= V(z, q) - C(z, p) + rW(z, q), \quad (3)$$

$$E_r(z, p, q) = L_r + rL_s$$
$$= W(z, q) + r(V(z, p) - C(z, p)). \quad (4)$$

The next step is to specify the pure strategies available to the players. The signaller can choose its strategy as a function of the original "move-by-nature", i.e. its state, whereas the receiver's strategy is a function of the signaller's behaviour. Therefore, four pure strategies are available to both the signaller and the receiver: (1) signaller (what to do if High, what to do if Low): (S, S), $(S, N)(N, S)$ and (N, N); (2) receiver (what to do if the signaller signals, what to do if not): (U, U),(U, D),(D, U) and (D, D). However, communication occurs in only four of the strategy pairs, when the signaller provides a variable signal and the receiver responds selectively to it. These four cases are formally identical, which means that the same analysis can be carried out for each pair, thus, it is enough to analyse only one of them $[(S, N)(U, D)]$ (Hurd, 1995).

3. Conditions for the Evolutionarily Stable Signalling under Conflict of Interest

In order for communication to occur, two conditions must be fulfilled (Hurd, 1995). The first one concerns the receiver and states that it should respond differently to the different states of the signaller, accepting a High state one with an Up reply and rejecting a Low state one with Down. That is, the following condition must hold (receiver's condition for communication—RCC):

$$E_r(H, S, U) > E_r(H, S, D),$$

$$E_r(L, S, U) < E_r(L, S, D),$$

that is:

$$W(H, U) + r(V(H, U) - C(H, S)) > W(H, D), + r(V(H, D) - C(H, S)), \quad \text{(RCCa)}$$

$$W(L, U) + r(V(L, U) - C(L, S)) < W(L, D) + r(V(L, D) - C(L, S)). \quad \text{(RCCb)}$$

The second condition states that the signaller should send different signals as a function of its state. If we assume that it should Signal when it is in a High state and it should Not signal when it is in a Low, then it gives the following condition (signaller's condition for communication—SCC):

$$E_s(H, S, U) > E_s(H, N, D)$$

$$E_s(L, S, U) < E_s(L, N, D)$$

that is,

$$V(H, U) - C(H, S) + rW(H, U) > V(H, D) - C(H, N) + rW(H, D), \quad \text{(SCCa)}$$

$$V(L, U) - C(L, S) + rW(L, U) < V(L, D) - C(L, N) + rW(L, D). \quad \text{(SCCb)}$$

The next step is to decide what relation holds between the signaller and the receiver concerning their final interest in the outcome of the game. Table 1 summarizes all the possibilities. The Up answer of the receiver denotes whether it is worthwhile for the receiver to turn up the signaller depending on the signaller's state (High, Low). The Up answer for the signaller denotes whether

TABLE 1

The relation between the signaller and the receiver concerning their final interest and its consequences on the signalling game

	Receiver's interest		Signaller's interest		Conflict of interest		Comment
Signaller's state	High	Low	High	Low	High	Low	
1			Up	Up	No	No	Always Up
2	Up	Up	Up	Down	No	Yes	
3			Down	Up	Yes	No	
4			Down	Down	Yes	Yes	
5			Up	Up	No	Yes	*
6	Up	Down	Up	Down	No	No	*
7			Down	Up	Yes	Yes	Always
8			Down	Down	Yes	No	Down
9			Up	Up	Yes	No	
10	Down	Up	Up	Down	Yes	Yes	Always D
11			Down	Up	No	No	
12			Down	Down	No	Yes	Always D
13			Up	Up	Yes	Yes	Always Down
14	Down	Down	Up	Down	Yes	No	
15			Down	Up	No	Yes	
16			Down	Down	No	No	

Note: Up and Down answers denote whether it is worth to turn Up or to accept the Up reply for the receiver and for the signaller respectively, depending on the state of the signaller (H,L). If both answer Up or both answer Down (given a certain state) then there is no conflict of interest between them.
* Indicate the two cases analysed by Maynard Smith.

it is worthwhile for the signaller to accept the receiver's Up reply (depending on the signaller's state). If both answer Up or both answer Down (given a certain state) then there is no conflict of interest between them. However, if one answers Up and the other Down then their interest is in conflict.

(a) Now, in the first four cases it is always worthwhile for the receiver to accept the signaller re*gardless whether it signals or not*. Obviously, no communication will evolve, the receiver will always respond with an Up reply.

(b) No communication evolves in the last four cases just for the opposite reason. Since it never pays the receiver to respond, it will always turn the signaller Down.

(c) The second and the third groups of four cases are formally identical hence it is enough to analyse only one of them. Let us continue with the second four. In two of these cases, in the seventh and eighth, there is no communication again for the following reasons, respectively: either their interests are always in conflict, or it is never worthwhile for the signaller to accept the receiver's Up reply. Finally, two cases remain: the fifth and the sixth. In the sixth case there is no conflict of interest between the players, and as Maynard Smith has pointed out (1991), cost-free communication can be an ESS and can maintain honest signalling. The most interesting case is of the fifth when there is a conflict of interest between the two participants. Much of the debate in the literature concerns this situation (Grafen, 1990; Maynard Smith, 1991, 1994; Hurd, 1995). In this case it is always beneficial for the signaller to accept the receiver Up reply whereas it pays for the receiver only to give Up reply if the signaller is in a High state. Since the second part of this statement is included in the RCC, it creates the following constraint for the signaller (conflict of interest condition—CIC):

$$E_s(H,*,U) > E_s(H,*,D),$$

$$E_s(L,*,U) > E_s(L,*,D),$$

that is,

$$V(H,U) + rW(H,U) > V(H,D) + rW(H,D) \quad \text{(CICa)}$$

$$V(L,U) + rW(L,U) > V(L,D) + rW(L,D) \quad \text{(CICb)}$$

where $*$ denotes "wildcard" (S or N). Finally, the cost of signalling must be defined. This can be different for High- and Low-state signallers:

$$C_h = C(H,S) - C(H,N) \quad (5)$$

$$C_l = C(L,S) - C(L,N), \quad (6)$$

where C_h and C_l denote the difference in cost between the two possible behaviours (Signal, Not to signal) for High- and Low-state individuals, respectively.

Similarly the receiver's fitness and the value of the receiver's reply can be defined depending on the signaller's state (H, L):

$$W_h = W(H,U) - W(H,D), \quad (7)$$

$$W_l = W(L,U) - W(L,D) \quad (8)$$

$$V_h = V(H,U) - V(H,D) \quad (9)$$

$$V_l = V(L,U) - V(L,D). \quad (10)$$

In conclusion it can be seen that evolutionary stable signalling under conflict of interest gives three conditions to be met (RCC, SCC, and CIC). In the following section the consequences of these conditions are analysed concerning the cost of signalling (C_h, C_l), in connection with Grafen's result (1990). Since it is possible that the value of the receiver's response is the same for all kinds of signallers ($V_h = V_l$), and the cost of the signal may not depend necessarily on the signaller's state ($C_h = C_l$), this gives rise to four possible combinations. If r can equal zero then this doubles the figure. The receiver's fitness may not depend on the signaller's state as well ($W_h = W_l$), in this case r must be greater than zero otherwise RCC cannot be fulfilled. The case where $W_h = W_l$, $V_h \neq V_l$, $C_h = C_l$ and $r > 0$ was analysed by Maynard Smith as the Philip Sidney game, and the case where $W_h \neq W_l$, $V_h = V_l$,

$C_h \neq C_l$ and $r = 0$ by Hurd in his Basic Action–Response Game. It is clear that most of the cases have not been analysed yet.

4. Results

In his paper Grafen has derived three "main handicap results" concerning evolutionarily stable signalling:

(a) signalling is honest, i.e. signals reveal the true quality of the signaller,
(b) signals are costly to produce,
(c) signals are more costly for worse signallers.

In the terminology of the present model these mean:

(a) that signallers in different states (High, Low) use different signals (Signal, Not to signal),
(b) that both C_h and C_l are greater than zero,
(c) C_l is greater than C_h.

Grafen's result (a) tells us what is meant by honest signalling and it is synonymous with SCC. Thus the question remains whether Grafen's results (b) and (c) must hold under the three conditions listed in the previous section (RCC, SCC, CIC).

The most general case is when the signaller's fitness depends on the signaller's state and both player's fitness' are influenced by the survival of the other. Then RCC, SCC and CIC have the form as it was given in the previous section. Rearranging them gives the following inequalities:

$$W_h + rV_h > 0, \quad \text{(IRa)}$$

$$W_l + rV_l < 0, \quad \text{(IRb)}$$

$$V_h + rW_h > C_h, \quad \text{(ISa)}$$

$$V_l + rW_l < C_l \quad \text{(ISb)}$$

$$V_h + rW_h > 0, \quad \text{(ICa)}$$

$$V_l + rW_l > 0. \quad \text{(ICb)}$$

Figure 1 shows the resulting three regions (1–3) where honesty is evolutionarily stable as a

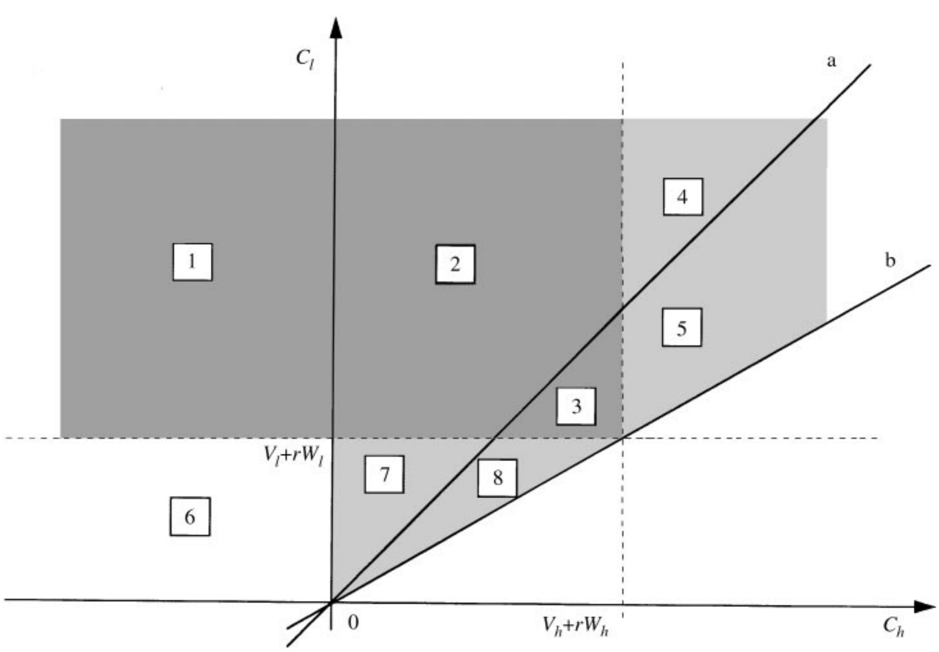

FIG. 1. The first case where $C_h \neq C_l$, $V_h \neq V_l$ and $r > 0$. On line a, $C_h = C_l$. On line b, the $C_h/(V_h + rW_h) = C_l/(V_l + rW_l)$ equation holds. There are three regions (1–3) where honest communication is stable (dark grey). In region 1 C_h is smaller than zero, thus Grafen's (b) does not hold. In region 2 both Grafen's (b) and (c) hold. In region 3 $C_h > C_l$, thus Grafen's (c) holds only if it is interpreted in terms of relative costs. Note regions 4, 5, 7, and 8 (light grey) where both Grafen's (b) and (c) hold yet dishonesty or no signalling is evolutionarily stable.

function of C_h and C_l. In the first region C_h is smaller than zero, thus Grafen's (b) does not hold. In the second region both Grafen's (b) and (c) hold. Finally, in the third region $C_h > C_l$, thus Grafen's (c) holds only if it is interpreted in terms of relative costs. From the figure and the inequalities we can see:

(1) that Grafen's (b) is not necessarily fulfilled, since the only constraint upon C_h is $V_h + rW_h > C_h$, that is, C_h can be negative;
(2) that Grafen's (c) is also not a necessary condition for ESS communication if it is interpreted in terms of simple costs since $V_h + rW_h$ can be greater than $V_l + rW_l$, thus the condition $C_h > C_l$ can be met which contradicts Grafen's (c). However, in terms of relative costs, when the cost of the signal divided by the overall benefit, i.e. $C_h/(V_h + rW_h)$ and $C_l/(V_l + rW_l)$, condition (c) must hold since for a high-quality signaller the numerator is always greater than the denominator whereas for a low-quality signaller just the opposite is the case (IS). Thus the relative cost is always greater for a low quality individual than for a high one. However, as can be seen from Fig. 1, even if Grafen's (c) is interpreted in terms of relative costs, it is still not a sufficient condition for ESS signalling. There are four regions (4, 5, 7, and 8) where the notion of the differential cost holds but honest communication is not evolutionarily stable. In the fourth and fifth case this is so because it is not worth even for high-quality individuals to signal, and in the seventh and eighth case for the opposite reason, that is, it is worth even for low-quality individuals to signal.

The same analysis can be carried out if we assume that $V_h = V_l$, or $C_h = C_l$, or $r = 0$ or any combination of these. The assumption $W_h = W_l$ has no qualitative effect on the results except that it implies $r > 0$ because if both $W_h = W_l$ and $r = 0$ then RCC cannot be met hence no communication can evolve. For instance, let us analyse the case: $V_h = V_l$, $C_h \neq C_l$, $r > 0$. This gives the following inequalities for RCC, SCC and CIC:

$$W(H,U) + r(V(U) - C(H,S)) > W(H,D) + r(V(D) - C(H,S)), \quad \text{(RCCa)}$$

$$W(L,U) + r(V(U) - C(L,S)) < W(L,D) + r(V(D) - C(L,S)), \quad \text{(RCCb)}$$

$$V(U) - C(H,S) + rW(H,U) > V(D) - C(H,N) + rW(H,D), \quad \text{(SCCa)}$$

$$V(U) - C(L,S) + rW(L,U) < V(D) - C(L,N) + rW(L,D), \quad \text{(SCCb)}$$

$$V(U) + rW(H,U) > V(D) + rW(H,D), \quad \text{(CICa)}$$

$$V(U) + rW(L,U) > V(D) + rW(L,D). \quad \text{(CICb)}$$

After rearrangement we have

$$W_h + rV > 0 \quad \text{(IRa)}$$

$$W_l + rV < 0, \quad \text{(IRb)}$$

$$V + rW_h > C_h, \quad \text{(ISa)}$$

$$V + rW_l < C_l, \quad \text{(ISb)}$$

$$V + rW_h > 0, \quad \text{(ICa)}$$

$$V + rW_l > 0. \quad \text{(ICb)}$$

The inequalities resulting from RCC, SCC and CIC for all the eight possible cases are summarized in Table 2 (in each case it is assumed that $W_h \neq W_l$). Note, that in the last case where the benefit (V) and the cost (C) are the same for high- and low-quality individuals and $r = 0$, evolutionarily stable signalling is not possible. The table also shows in each case the consequences on Grafen's results. As can be seen, Grafen's (b) holds only if we do not allow the cost of the signal to vary with the signaller's state, that is, we assume that the cost is the same for all individuals ($C_h = C_l$). This follows from the fact that both the high- and the low-quality signaller's expected fitness is always greater than zero (see Table 2,

TABLE 2

IR, IS and IC for the eight possible combinations of $V_h = V_l$, $C_h = C_l$, $r = 0$, and their consequences on Grafen's main handicap results

		IR	IS	IC	Grafen's (b)	Grafen's (c)	ROD
1	$V_h \neq V_l$ $C_h \neq C_l$ $r > 0$	$W_h + rV_h > 0$ $W_l + rV_l < 0$	$V_h + rW_h > C_h$ $V_l + rW_l < C_l$	$V_h + rW_h > 0$ $V_l + rW_l > 0$	No	Yes/no	Yes
2	$V_h = V_l$ $C_h \neq C_l$ $r > 0$	$W_h + rV_h > 0$ $W_l + rV_l < 0$	$V + rW_h > C_h$ $V + rW_l < C_l$	$V + rW_h > 0$ $V + rW_l > 0$	No	Yes/no	Yes
3	$V_h \neq V_l$ $C_h = C_l$ $r > 0$	$W_h + rV_h > 0$ $W_l + rV_l < 0$	$V_h + rW_h > C$ $V_l + rW_l < C$	$V_h + rW_h > 0$ $V_l + rW_l > 0$	Yes	Yes/no	No
4	$V_h = V_l$ $C_h = C_l$ $r > 0$	$W_h + rV > 0$ $W_l + rV < 0$	$V + rW_h > C$ $V + rW_l < C$	$V + rW_h > 0$ $V + rW_l > 0$	Yes	Yes/no	No
5	$V_h \neq V_l$ $C_h \neq C_l$ $r = 0$	$W_h > 0$ $W_l < 0$	$V_h > C_h$ $V_l < C_l$	$V_h > 0$ $V_l > 0$	No	Yes/no	Yes
6	$V_h = V_l$ $C_h \neq C_l$ $r = 0$	$W_h > 0$ $W_l < 0$	$V > C_h$ $V < C_l$	$V > 0$	No	Yes/yes	yes
7	$V_h \neq V_l$ $C_h = C_l$ $r = 0$	$W_h > 0$ $W_l < 0$	$V_h > C$ $V_l < C$	$V_h > 0$ $V_l > 0$	Yes	Yes/no	No
8	$V_h = V_l$ $C_h = C_l$ $r = 0$	$W_h > 0$ $W_l < 0$	$V > C$ $V < C$	$V > 0$	It is not possible		

Note: At Grafen's (c) the two different interpretations are: relative cost/simple cost. The last column informs whether regions of dishonesty (ROD) can be found in the face of Grafen's (b) and (c) conditions (for examples see Fig. 1.). The third case was analysed by Maynard Smith and the sixth by Hurd.

second column: IC); and that the cost for high-quality signallers should be smaller than this (creating an upper bound for C_h but not a lower one) but the cost for low-quality signallers should be greater than this (creating a lower bound for C_l but not a upper one) (see the conditions for IS, Table 2, first column). However, the lower and the upper bound combine together if we do not allow the cost to vary independently with the signaller's state. Thus, in this case the cost should be always greater than zero (see Fig. 2 for an example). Grafen's (c) holds in each case provided it is interpreted in terms of relative costs. However, if it is interpreted in terms of simple costs then it holds only in the sixth case because in this case the expected fitness' are the same for high- and low-quality signallers but the cost can vary.

FIG. 2. The third case analysed by Maynard Smith as the Sir Philip Sidney game. The thick line denotes the region where honest communication is stable. $C_h = C_l$ thus Grafen's (c) can be valid only if it is interpreted in terms of relative costs. Note that only in one dimension is it true that Grafen's (b) and (c) are necessary and sufficient conditions for evolutionarily stable reliable communication.

5. Discussion

The present model shows that in a simple discrete communication game two types of outcomes are possible, depending on whether the signaller's fitness and the cost of the signal depend on the signaller's state and on the assumption of inclusive fitness (or on other kinds of influence between the signaller's and the receiver's fitness). On the one hand, if the cost of the signal can vary with the signaller's state ($C_h \neq C_l$) Grafen's (b) is not a necessary condition for evolutionarily stable communication, that is, signals need not be costly. On the other hand ($C_h = C_l$), Grafen's (b) holds only because the cost of the signal is the same for all kind of individuals. In all cases Grafen's (c) holds if it is interpreted in terms of relative costs, i.e. the signal should be more costly for low-quality individuals not in itself but in relation to the whole benefit of the interaction. However, if it is interpreted in terms of simple costs then there is only one case in which Grafen's (c) holds, the sixth, and by chance exactly this case was analysed by Hurd. This makes clear why he has concluded, even though he was thinking in terms of simple costs, that: "Grafen's (c) follows directly from the definition of communication". Yaschi (1995) has also arrived at a similar conclusion, although he analysed a different case. He assumed that both the value of the signaller's response and the cost of the signal can vary with the signaller's state ($C_h \neq C_l$, $V_h \neq V_l$, $r = 0$). As we know, in this case Grafen's (c) holds only in terms of relative costs. Yaschi, however, made an additional assumption, that the value of the receiver's response should be always greater for low-quality individuals than that for high-quality ones ($V_h < V_l$). This explains why he could assert that his results are consistent with Grafen's result even in terms of simple cost. However, even if Grafen's (c) is interpreted in terms of relative costs in a discrete communication game it is not enough in itself to guarantee stable and reliable communication (see Fig. 1). The notion of the differential cost is a necessary but not a sufficient condition.

The present model shows that the stability conditions for honest and evolutionarily stable communication are as follows: (a) it is beneficial to a high-quality individual to use the signal (i.e. $V_h + rW_h > C_h$), and (b) a low-quality individual cannot expect an overall benefit from using it (i.e. $V_l + rW_l < C_l$). Note, that the cost of the signal for a low-quality individual (C_l) should exceed the difference between the benefits of the possible alternative outcomes $[V(L,U) + rW(L,U) - [V(L,D) - rW(L,D) = V_l + rW_l$, see eqn (SCCb)] and not the benefit resulting from the use of the signal ($V(L,U) + rW(L,U)$). That is, both the signal cost and the possible benefits should be treated as relative quantities (for a more detailed discussion of signal cost, see Bergstrom & Lachmann, 1998). In this sense these conditions are identical to Zahavi's (1993) formulation in which he says that: ". . . the investment in a signal cannot be greater than the potential gain from using it properly. A signal is reliable when the investment required for its use is greater than the potential gain a cheater would make from using it improperly. The investment should be acceptable to an honest signaller and prohibitive to a cheater".

The model shows that Grafen's results, which were derived from a continuous model cannot be "translated" directly to a discrete case. That is, the handicap principle may have to be formulated differently in case of discrete and in case of continuous models (for a detailed discussion of the continuous case, see Getty, 1998a, b). Only in one dimension (regarding C, that is, when $C_h = C_l$) is it true that Grafen's (b) and (c) are necessary and sufficient conditions for ESS signalling. In two dimensions (i.e. $C_h \neq C_l$), although there is an overlap between the regions where Grafen's conditions are valid and the regions where honest communication is evolutionarily stable, the overlap is far from complete. To present the difference more clearly it is useful to compare them graphically. Figure 3 depicts (a) Grafen's definition in terms of simple costs, (b) Grafen's definition in terms of relative costs, and (c) the stability conditions for honest and evolutionarily stable communication. Shaded areas represent the honest and evolutionarily stable regions according to the appropriate definition. Note, that neither Grafen's definition is a subset of the stability conditions nor the stability conditions are subsets of Grafen's definition. Moreover, Fig. 1 reveals that in the regions 4, 5, 7 and 8 both Grafen's (b) and (c) hold (in terms of

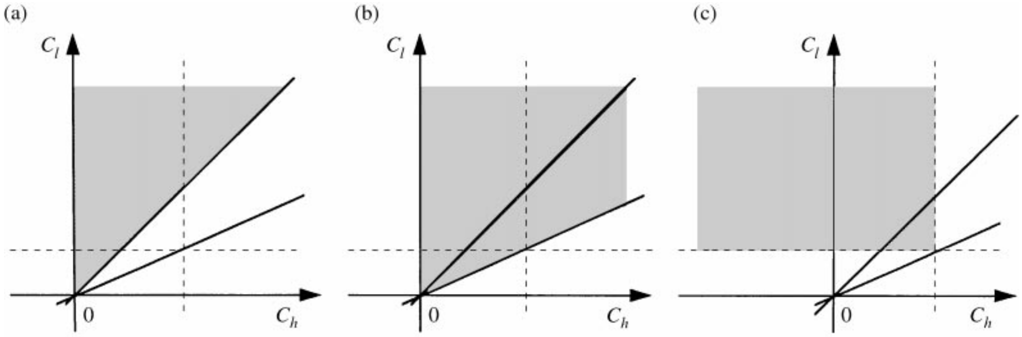

FIG. 3. Comparison of the two different definitions of honest communication. (a) Grafen's definition in terms of simple costs, (b) Grafen's definition in terms of relative costs, and (c) the stability conditions defined by IR, IS and IC. Shaded areas represent the honest and evolutionarily stable regions predicted by the appropriate definition.

relative costs), yet in the regions 4 and 5 signalling will not evolve because it is not worth to signal even for high-quality signallers, and in the regions 7 and 8 deception is evolutionarily stable because it is worth to signal even for a low-quality signaller [in face of Grafen's (b) and (c)!]. Such regions can be found wherever C_h differs from C_l, that is, in four cases out of the possible eight. On the other hand, in region 1 Grafen's (b) does not hold, yet honest communication is evolutionarily stable. Finally, it should be emphasized that the fact that Grafen's definition cannot be "translated" directly to discrete models does not question its validity in the continuous case. Moreover, the present model is fully in line with the recent result of Getty (1998a, b) which says that in an evolutionarily stable signalling system high-quality individuals should convert advertisement into fitness more efficiently than low-quality individuals do.

The model also highlights the difference in the assumptions of two previous models concerning simple communication games (Maynard Smith, 1991; Hurd, 1995). The crucial one is that while in Hurd's model the cost of a signal can vary with the signaller's state, in Maynard Smith's model it cannot. For this reason Maynard Smith's choice may not be the best one, since his model is to emphasize that signalling must be costly but if his condition is rejected it turns out to be otherwise. However, Maynard Smith points out a very important conclusion, that cost-free communication can be stable if there is no conflict of interest between the participants.

I thank John Maynard Smith, Eörs Szathmáry and István Scheuring for helpful comments and discussion and Amotz Zahavi and an anonymous referee for helpful comments. This work was supported by the Pál Juhász-Nagy junior-fellowship of Collegium Budapest.

REFERENCES

BERGSTROM, C. T. & LACHMANN, M. (1998). Signaling among relatives. III. Talk is cheap. *Proc. Nat. Acad. Sci. U.S.A.* **95,** 5100–5105.

DAWKINS, M. S. & GUILFORD, T. (1991). The corruption of honest signalling. *Anim. Behav.* **41,** 865–873.

GETTY, T. (1998a). Handicap signalling: when fecundity and viability do not add up. *Anim. Behav.* **56,** 127–130.

GETTY, T. (1998b). Reliable signalling need not de a handicap. *Anim. Behav.* **56,** 253–255.

GRAFEN, A. (1990). Biological signals as handicaps. *J. theor. Biol.* **144,** 517–546.

HURD, P. L. (1995). Communication in discrete action–response games. *J. theor. Biol.* **174,** 217–222.

HURD, P. L. (1997). Is signalling of fighting ability costlier for weaker individuals? *J. theor. Biol.* **184,** 83–88.

KREBS, J. R. & DAWKINS, R. (1984). Animal signals: mindreading and manipulation. In: *Behavioural Ecology. An Evolutionary Approach* (Krebs, J. R. and Davies, N. B., eds) pp. 380–402. Oxford: Blackwell Scientific Publications.

MAYNARD SMITH, J. (1982). *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.

MAYNARD SMITH, J. (1991). Honest signalling: the Philip Sidney game. *Anim. Behav.* **42,** 1034–1035.

MAYNARD SMITH, J. (1994). Must reliable signals always be costly? *Anim. Behav.* **47,** 1115–1120.

MAYNARD SMITH, J. & HARPER, D. G. C. (1995). Animal signals: models and terminology. *J. theor. Biol.* **177,** 305–311.

MAYNARD SMITH, J. & PRICE, G. R. (1973). The logic of animal conflict. *Nature* **246,** 15–18.

NUR, N. & HASSEN, O. (1984). Phenotypic plasticity and the handicap principle. *J. theor. Biol.* **110,** 275–297.

POMIANKOWSKI, A. (1987). Sexual selection: the handicap principle does work—sometimes. *Proc. R. Soc. Lond. B* **231,** 123–145.
YACHI, S. (1995). How can honest signalling evolve? The role of handicap principle. *Proc. R. Soc. Lond. B* **262,** 283–288.
ZAHAVI, A. (1975). Mate selection—a selection for handicap. *J. theor. Biol.* **53,** 205–214.
ZAHAVI, A. (1977). The cost of honesty. *J. Theor. Biol.* **67,** 603–605.
ZAHAVI, A. (1993). The fallacy of conventional signalling. *Phil. Trans. R. Soc. Lond. B* **340,** 227–230.