Is Signalling of Fighting Ability Costlier for Weaker Individuals?

Peter L. Hurd†

Division of Ethology, Department of Zoology, University of Stockholm, S-106 91 Stockholm, Sweden

Using a simple model of signaling of fighting ability, I demonstrate that; (1) conventional, cost-free, signals of fighting ability can be an ESS, (2) signals with significant costs can be used at ESS as long as they are used to indicate weakness rather than strength, (3) that if a set of signals is used to indicate a set of fighting abilities through their costs, they must decrease in cost for stronger signalers. The reason for this is that individuals of higher fighting ability have less to gain by avoiding escalated contests, and are thus more sensitive to signal costs. These results are of particular relevance to badges of status and other simultaneous signals used to settle contests over minor resources.

© 1997 Academic Press Limited

1. Introduction

Conflicts over resources are ultimately resolved through fighting. Two factors determine the outcome of a fight, the ability to inflict cost and the transfer of information. Physical fighting is costly and risky (Geist, 1974) and communication seems a good way to reduce or limit these costs (Tinbergen, 1951; Parker, 1974; Rohwer, 1982). Signals of fighting ability are used to settle minor contests in some species (Ketterson, 1979; Fugel & Rothstein 1987; Møller, 1987). Two general types of signals may be employed in agonistic communication, performance signals, and strategic signals. Performance signals have also been termed unambiguous signalling (Maynard Smith, 1982), assessment signalling (Maynard Smith & Harper, 1988), and condition dependent handicaps (Grafen, 1990). In performance signalling some signallers may perform signals that others cannot. Strategic signals are those which all signalers are capable of making. Strategic signalling seems vulnerable to cheating as variation in the choice of behaviour is maintained only due to differing cost and benefit factors between signals and signallers.

†E-mail: pete@zool.su.se

All signals are stabilized by costs. If there is to be more than one optimal behaviour, then the costs and benefits of the various options must be different. Early game theorists understood that some sort of cost had to balance, and exceed, the gains possible though "bluffing" (Maynard Smith, 1974, 1982) but were not able to provide a reasonable source of this cost. Grafen (1990) provided a model in which the wastefullnes of a signal, its "handicap", stabilized signalling.

The costs which stabilize strategic signalling of fighting ability are hypothesized to have one of three sources. Firstly, signals of high fighting ability may be so costly to produce that weaker individuals cannot profit from their use (Fugel & Rothstein, 1987; Møller, 1987; Grafen, 1990; Slotow *et al.*, 1993). Secondly, the properties of the signal may interact with the receiver's reaction to produce costs unprofitably high for some signallers (Enquist *et al.*, 1985; Adams & Mesterton-Gibbons, 1995). Thirdly, social interactions may somehow enforce the convention linking signal form and meaning (sensu Smith, 1977), and it is the high cost of the receiver's reaction to an inflated signal that inferior signallers cannot afford (Rohwer & Ewald, 1981). In the first case are handicapped signals (Zahavi, 1975, 1977; Grafen,

1990; Hurd, 1995), in the second case are risks (Enquist *et al.*, 1985) or vulnerability handicaps (Adams & Mesterton-Gibbons, 1995) and the third case includes conventional signals (Enquist, 1985; Guilford & Dawkins, 1995). Handicapped signals are maintained by costs which are purely a function of the signal and signaller state (Hurd, 1995) (or signal, state, and reply in the case of risks or "vulnerability handicaps"), whereas conventional signals are stabilized by costs which are purely a function of the receiver's reply and signaller state (Enquist, 1985).

Using a simple model of conventional signalling [based on Enquist's (1985) threat model] I will demonstrate that the effect of adding an arbitrary cost to a conventional signal fundamentally alters the use of that signal in precisely the opposite manner that the handicap principle suggests that it should. I suggest that signals indicating fighting ability will either be conventional, or more costly for weaker signallers.

2. The Conventional Signalling Game

This game is based on Enquist's (1985) Example 1. Two individuals compete over some non-divisible resource. These individuals differ in fighting ability such that stronger individuals always prevail in physical contests (for simplicity's sake we shall assume that all strength classes are equally represented in the population). Contestants know their own fighting ability, but not that of their opponent.

Each contestant makes a simultaneous display (A and B are alternative signals), and observes the opponent's display before choosing a behaviour (this models situations such as plumage badges, where the signal is effectively simultaneous). The possible behaviours are, Attack, Pause-Attack, and Flee. A Pause-Attacking individual attacks after a pause allowing a fleeing opponent to escape without physical conflict, but remains to fight against attacking or pause-attacking opponent.

The basic pay-offs are presented in Table 1. The pay-offs are essentially those of Enquist (1985, un-numbered table on page 1156). V is the value of the resource to the winner, this is presumed to be equal for both contestants. $C(a)$ is the cost of physical combat against an opponent who is a states stronger (negative a values signify a weaker opponent), and F_A is the cost of Attacking an opponent who is Fleeing. There is also assumed to be some cost, F_P, in pausing before attacking an opponent who is attacking directly.

Being attacked is assumed to involve a significant cost, or expected cost, regardless of the decision to fight or flee. These costs are of the same order as V if the opponent is of equal or greater strength. (Contests over more valuable resources are expected to involve greater total costs, but they are unlikely to accumulate within the time-span of a single decision point. Contests over less valuable resources, such as a few seeds at a bird feeder, are constantly at the borderline of net gain and net loss. In such contests, a single fast peck from the opponent is more likely to make a huge difference in the relative value of contesting the resource, and to occur effectively instantly.)

It is assumed that $C(a)$ is an increasing function of a, and $C(a)$, V, F_P and F_A are all positive, [although not a necessary assumption, it is biologically realistic to assume that $C(a) > F_P$ and F_A]. The $\frac{1}{2}V$ pay-off to players who both flee is traditional in models of aggression (Maynard Smith, 1982) if the pay-off to this eventuality were nil, no behaviours would change and all results would be the same. The pay-offs are entirely independent of the actual display used,

TABLE 1
Pay-offs as a function of State of Behaviour

Ego strength	Opponent behaviour	Ego behaviour					
		Opponent strong			Opponent weak		
		Attack	Pause-Attack	Flee	Attack	Pause-Attack	Flee
Strong	Attack	$\frac{1}{2}V - C(0)$ [1]	$\frac{1}{2}V - C(0) - F_P$	$-C(0)$	$V - C(-1)$ [9]	$V - C(-1) - F_P$	$-C(-1)$ [6]
	Pause-Attack	$\frac{1}{2}V - C(0)$ [8]	$\frac{1}{2}V - C(0)$	0 [5]	$V - C(-1)$	$V - C(-1)$	0
	Flee	$V - F_A$	V	$\frac{1}{2}V$	$V - F_A$	V [2]	$\frac{1}{2}V$
Weak	Attack	$-C(1)$	$-C(1) - F_P$	$-C(1)$ [10]	$\frac{1}{2}V - C(0)$ [4]	$\frac{1}{2}V - C(0) - F_P$	$-C(0)$ [7]
	Pause-Attack	$-C(1)$	$-C(1)$	0 [3]	$\frac{1}{2}V - C(0)$	$\frac{1}{2}V - C(0)$	0
	Flee	$V - F_A$	V [11]	$\frac{1}{2}V$	$V - F_A$	V	$\frac{1}{2}V$

F_P is the cost of Pausing before Attacking an Attacker, and F_A is the cost of Attacking a Fleeing opponent. $C(0)$, corresponds to Enquists's C, $C(1)$ with Enquist's D, and $C(-1)$ with Enquist's " $<$ " sign in the $<V$ pay-offs to Strong players showing B against opponents who also show B.

TABLE 2
Pay-offs against a conventional signalling population

(a) Strong

Display	Opponent display	Opponent behaviour	Behaviour		
			Attack	Pause-Attack	Flee
A	A	Attack	$\frac{1}{2}V - C(0)$ [1]	$\frac{1}{2}V - C(0) - F_P$	$-C(0)$
	B	Flee	$V - F_A$	V [2]	$\frac{1}{2}V$
B	A	Pause-Attack	$\frac{1}{2}V - C(0)$ [8]	$\frac{1}{2}V - C(0)$	0 [5]
	B	Attack	$V - C(-1)$ [9]	$V - C(-1) - F_P$	$-C(-1)$ [6]

(b) Weak

Display	Opponent display	Opponent behaviour	Behaviour		
			Attack	Pause-Attack	Flee
A	A	Attack	$-C(1)$	$-C(1) - F_P$	$-C(1)$ [10]
	B	Flee	$V - F_A$	V [11]	$\frac{1}{2}V$
B	A	Pause-Attack	$-C(1)$	$-C(1)$	0 [3]
	B	Attack	$\frac{1}{2}V - C(0)$ [4]	$\frac{1}{2}V - C(0) - F_P$	$-C(0)$ [7]

Pay-offs to different behaviours as a function of strength and display, assuming the opponent is using the ESS strategy. Opponents showing Display A are strong, and those showing Display B are weak. Numbers in square brackets refer to cells discussed in the text.

allowing truly conventional signalling (Enquist, 1985).

These pay-offs differ from Enquist's original parameters in that the F_P and F_A parameters implied by " $<$ " symbols have been made explicit. In addition, $a - C(0)$ has been added so to the cell [7] so that the Strong vs. Strong and Weak vs. Weak pay-offs are the same (See Table A1 for a generic expression of pay-offs).

We shall determine the conditions under which the pure communication strategy S_{AB}:

"Use display A when strong and B when weak, then Attack opponents who signal an equal strength, Flee from opponents who signal a stronger state, and Pause-Attack opponents signalling a weaker state"

is an ESS. Of course a mirror strategy, S_{BA}, exists in which the use of A and B are reversed, but this can be (temporarily) ignored without loss of generality. Pay-offs to a player playing against a population of S_{AB} strategists are shown in Table 2 (Note that the rows in Table 1 which are not used by a population playing S_{AB} do not appear in Table 2). An individual playing S_{AB} receives the pay-offs numbered [1–4] in Table 2 (numbers in square brackets referr to pay-offs in Tables 1 & 2). The pay-offs to an S_{AB} strategist are null under the Weak vs. Strong condition and positive under the conditions: Strong vs. Strong, Strong vs. Weak and Weak vs. Weak, while

$$\frac{1}{2}V > C(0) \tag{1}$$

2.1. PAYOFFS TO ALTERNATIVE STRATEGIES

There are three relevant alternative strategies, "Coward", "Trojan", and "Bluff", (C, T and B respectively). These strategies are superior replies to S_{AB} under some parameter values. We shall proceed by deducing the parameter values for which S_{AB} is an ESS, [all other alternative strategies are inferior to C, T and B regardless of parameter values, given that F, $C(a)$ and V are positive].

2.1.1. Coward

The Coward strategy minimizes costs incurred by always signalling weakness, then fleeing from all opponents. A strong signaller may always signal weakness and flee with a pay-off of zero. For S_{AB} to be an ESS is must be the case that S_{AB} pays more than C, [1] + [2] > [5] + [6], which requires,

$$1\frac{1}{2}V > C(0) - C(-1). \tag{2}$$

A weak player may signal weakness, then flee, this is unprofitable when [4] > [7], which requires that,

$$V > 0. \tag{3}$$

2.1.2. Bluff

A bluffer follows the same behavioural rules as S_{AB} when strong, but when weak signals strength then flees if the opponent also signals strength, and pause-attacks if the opponent signals a weak state. A weak player bluffing strength receives pay-offs [10]

and [11] rather than [3] and [4], and is unprofitable when,

$$\tfrac{1}{2}V < C(1) - C(0). \qquad (4)$$

2.1.3. *Trojan Sparrow*

The "Trojan" strategy in which a strong individual signals weakness then proceeds to Attack his opponent pays [8] and [9] rather than [1] and [2]. This is unprofitable as long as

$$C(-1) > 0. \qquad (5)$$

When these conditions are met, S_{AB} and S_{BA} are the only ESS solutions, parameter values producing evolutionarily stable conventional signalling are presented in Fig. 1. We will now proceed by applying a handicapping cost to one of the displays.

3. Handicaps

How will this game change when we add an arbitrary cost, H, to the performance of the A display? Signals will remain conventional as long as $[1] + [2] - 2H > [8] + [9]$, which requires

$$H < \tfrac{1}{2}C(-1). \qquad (6)$$

While this condition is met both S_{AB} and S_{BA} are ESSes. If H is larger than this threshold value the Trojan strategy is more profitable. When this is so S_{AB} is no longer an ESS, though S_{BA} still is. Stronger

signallers will use the costless signal and the weaker players will switch to using the handicapped signal. This will remain a stable signalling ESS as long as Bluff remains unprofitable, $[3] + [4] - 2H > [10] + [11]$, which requires,

$$H < \frac{C(1) - C(0) - \tfrac{1}{2}V}{2}. \qquad (7)$$

So more costly signals can occur in a signalling ESS when the "weakness" signal is handicapped than when the "strength" signal is. This result holds true when there are N strength classes and signals (Appendix).

The net profitability of the paradoxical strategy S_{BA} is maintained, as long as

$$\tfrac{1}{2}V - 2H > C(0). \qquad (8)$$

While pay-off to a weak signaller opposing a stronger is negative, $-H$, it will still pay a weak individual to play the game, and use the handicapped signal when (8) is met.

Inequalities (6), (7) and (A.5) present the threshold values of costs that are actually paid at ESS. If signallers of different states differ in their capacity to pay costly signals (7) and (A.5) still apply. Though the costs of deviating from the evolutionarily stable strategy will no longer be accurately predicted by the model, it is still known that such deviations are not optimal, and the ESS derived above still holds.

4. Discussion

Most previous models of agonistic communication have modelled the signalling of aggressiveness (the tendency to escalate) stabilized by frequency dependent effects (Maynard Smith & Harper, 1988; Johnstone & Norris, 1993; Kim, 1995). In this case the benefit gained by a high status signal must be balanced by an equal cost (Maynard Smith, 1982; Järvi & Bakken, 1984; Studd & Robertson, 1985). The assumption that fighting ability and differences in relative resource value are insignificant makes the biological relevance of these models questionable.

Two analytic models claim to analyse signalling of fighting ability (Gardner & Morris, 1989; Adams & Mesterton-Gibbons, 1995). The Gardner and Morris model does not model signalling (Adams & Mesterton-Gibbons, 1995). The Adams & Mesterton-Gibbons (1995) model is an action-response game (Hurd, 1995) in which one individual signals with one of two alternatives, and the receiver responds with one of two replies. Like other action-response games (Cho & Kreps, 1987; Grafen, 1990; Maynard Smith, 1991, 1994; Godfray, 1991; Johnstone & Grafen,
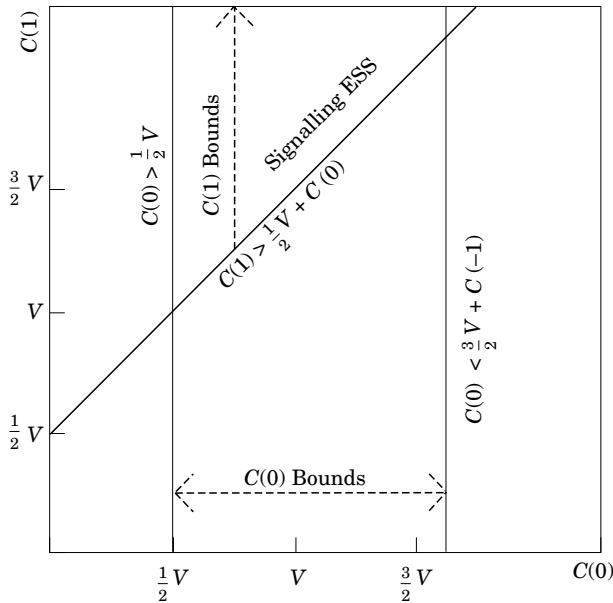

FIG. 1. A graphical representation of the parameter values for a communication ESS. The bounds on $C(0)$ are $C(0) > \tfrac{1}{2}V$ from inequality (1), $C(0) < \tfrac{3}{2}V + C(-1)$ from inequality (2), the bounds on $C(1)$, $C(1) > \tfrac{1}{2}V + C(0)$ is from inequality (4).

1993; Hurd, 1995), this model cannot be generalized to cover situations in which both individuals exchange information and base decisions upon each others' signal. Such exchanges of information typifies aggressive signalling. Action-Response games are traditionally used to model handicaps because they more resemble the mate quality signalling system, and they are strongly predisposed to producing handicapped results (Hurd & Enquist, in preparation).

Fights are assumed to end when assymmetries in fighting ability become apparent to the contestants, rather than when serious injury occurs [as in the Hawk-Dove game (Maynard-Smith, 1982)]. Empirical evidence (Enquist et al., 1988; Leimar et al., 1991; Englund & Olsson, 1990; Koops & Grant, 1993; Keeley & Grant, 1993) shows that the costs that are actually paid in resolving these contests are highest for closely matched opponents, which agrees with the results of this model.

Weak animals communicate their strength to avoid provoking stronger opponents, while strong individuals gain by avoiding unnecessary escalated contests. Since strength and weakness refer to fighting ability, the benefits of avoiding combat must be smaller for individuals who could win through force (as was noted by Adams & Mesterton-Gibbons, 1985), and thus the temptation to defect from the communication equilibrium is greater for stronger players. Adding an arbitrary cost to a threat display will affect the more cost sensitive individuals leading to the use of that handicapped display by weaker, not stronger, individuals (of course, if a costless third signal were added, the costly signal would go unused by any player with no detriment to stability). The cost paid by signallers at ESS will be either very small and irrelevant to the association of signals with state, or significant with the weaker signallers paying more. This effect is to be expected whenever resource value is equal, signalling is mutual and simultaneous (as it is in badge signals), and communicates fighting ability. State dependent ability to pay handicap costs may change the form of the expressed signal, but will not alter this effect.

REFERENCES

ADAMS, E. S. & MESTERTON-GIBBONS, M. (1995). The cost of threat displays and the stability of deceptive communication. *J. theor. Biol.* **175,** 405–421.

CHO, I. K. & KREPS, D. M. (1987). Signaling games and stable equilibria. *Q. J. Econ.* **102,** 179–221.

ENGLUND, G. & OLSSON, T. I. (1990). Fighting and assessment in the net-spinning caddis larva *Arctopsyche ladogensis*: a test of the sequential assessment game. *Anim. Behav.* **39,** 55–62.

ENQUIST, M. (1985). Communication during aggressive interactions with particular reference to variation in choice of behaviour. *Anim. Behav.* **33,** 1152–1161.

ENQUIST, M., PLANE, E. & ROED, J. (1985). Aggressive communication in fulmars (*Fulmarus glacialis*). *Anim. Behav.* **33,** 1007–1020.

FUGEL, G. N. & ROTHSTEIN, S. I. (1987). Experiments on the control of deceptive signals of status in white-crowned sparrows. *Auk* **104,** 188–197.

GARDNER, R. & MORRIS, M. R. (1989). The evolution of bluffing in animal contests: an ESS approach. *J. theor. Biol.* **137,** 235–243.

GEIST, V. (1974). On fighting strategies in animal combat. *Nature* **250,** 354.

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

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

GUILFORD, T. & DAWKINS, M. S. (1995). What are conventional signals? *Anim. Behav.* **49,** 1689–1695.

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

JÄRVI, T. & BAKKEN, M. (1984). The function of the variation in the breast stripe of the great tit (*Parus major*). *Anim. Behav.* **32,** 590–596.

JOHNSTONE, R. A. & GRAFEN, A. (1993). The continuous Sir Philip Sidney game: a simple model of animal signalling. *J. theor. Biol.* **156,** 215–234.

JOHNSTONE, R. A. & NORRIS, K. (1993). Badges of status and the cost of aggression. *Behav. Ecol. Sociobiol.* **32,** 127–134.

KEELEY, E. R. & GRANT, J. W. A. (1993). Visual information, resource value, and sequential assessment in convict cichlid (*Cichlasoma nigrofasciatum*) contests. *Behav. Ecol.* **4,** 345–349.

KETTERSON, E. D. (1979). Status signalling in dark-eyed juncos. *Auk* **96,** 94–99.

KIM, Y-G. (1995). Status signalling games in animal contests. *J. theor. Biol.* **176,** 221–231.

KOOPS, M. A. & GRANT, J. W. A. (1993). Weight asymmetry and sequential assessment in convict cichlid contests. *Can. J. Zool.* **71,** 475–479.

LEIMAR, O., AUSTAD, S. & ENQUIST, M. (1991). A test of the sequential assessment game: fighting in the bowl and doily spider *Frontinella pyramitela*. *Evolution* **45,** 862–874.

MAYNARD-SMITH, J. (1974). Theory of games and the evolution of animal contests. *J. theor. Biol.* **47,** 209–221.

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

MAYNARD-SMITH, J. (1991). Honest signalling: the Philip Sidney game. *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. (1988). The evolution of aggression: can selection generate variability? *Phil. Trans. R. Soc. Lond.* B **319,** 557–570.

MØLLER, A. P. (1987). Social control of deception among status signalling house sparrows *Passer domesticus*. *Behav. Ecol. Sociobiol.* **20,** 307–311.

PARKER, G. A. (1974). Assessment strategy and the evolution of animal conflicts. *J. theor. Biol.* **47,** 223–243.

ROHWER, S. (1982). The evolution of reliable and unreliable badges of fighting ability. *Am. Zool.* **22,** 531–546.

ROHWER, S. & EWALD, P. W. (1981). The cost of dominance and advantage of subordination in a badge signaling system. *Evolution* **32,** 670–673.

SLOTOW, R., ALCOCK, J. & ROTHSTEIN, S. I. (1993). Social status signalling in white-crowned sparrows: an experimental test of the social control hypothesis. *Anim. Behav.* **46,** 977–989.

SMITH, W. J. (1977). *The Behaviour of Communicating: an Ethological Approach.* Harvard University Press.

STUDD, M. V. & ROBERTSON, R. J. (1985). Evidence for reliable badges of status in territorial yellow warblers (*Dendroica petechia*). *Anim. Behav.* **33,** 1102–1113.

TINBERGEN, M. (1951). *The Study of Instinct.* Oxford: Clarendon.

ZAHAVI, A. (1975). Mate selection—a selection for a handicap. *J. theor. Biol.* **53,** 205–214.

ZAHAVI, A. (1977). The cost of honesty (Further remarks on the handicap principle). *J. theor. Biol.* **67,** 603–605.

APPENDIX

Multiple States

The pay-offs presented in Table 1 are presented in a generic form, describing pay-off to contestants in an N state version of this game, in Table A1.

We can easily extend the model to cover N states, we modify S to cope with an open ended set of states and signals,

"Use display A when state is N, B when state is $N-1$, C when state is $N-2$ ..., Attack opponents who signal an equal strength, Flee

TABLE A1
Pay-offs as a function of State and Behaviour

Opponent behaviour	Behaviour		
	Attack	Pause-Attack	Flee
Attack	$W - C(a)$	$W - C(a) - F_P$	$-C(a)$
Pause-Attack	$W - C(a)$	$W - C(a)$	0
Flee	$V - F_A$	V	$\frac{1}{2}V$

W is V if the opponent is weaker, $(a < 0)$, $\frac{1}{2}V$ is the opponent is of equal strength, $(a = 0)$, and 0 is the opponent is stronger, $(a > 0)$. F_P is the cost of Pausing before Attacking an Attacker, and F_A is the cost of Attacking a Fleeing opponent, $C(0)$, corresponds to Enquist's C, $C(1)$ with Enquist's D, and $C(-1)$ with Enquist's " $<$ " sign in the $<V$ pay-offs to Strong players showing B against opponents who also show B.

from opponents who signal a stronger state, and Pause-Attack opponents signalling a weaker state".

Bluffing is unprofitable when,

$$(N - \frac{2}{3})V < C(N - 1) - C(0) \qquad (A.1)$$

Conventional communication is the only pure ESS when this condition is met.

We can now add handicapping costs to these displays, H_s is the cost of the display associated with state s. At lower values of H, signals will remain interchangeable as long as S scores higher than all T.

$$w_S = (s - 1)V + \frac{1}{2}V - C(0) - NH_s \qquad (A.2)$$

$$w_T = (s - 1)V + \frac{1}{2}V - C(0)$$

$$- \sum_i^t C(-i) - NH_{s-t} \ t \in (1 \ldots s - 1) \quad (A.3)$$

where, t is the number of states lower than s that is signalled, which requires,

$$H_s - H_{s-t} < \frac{\sum_i^t C(-i)}{N} \ \forall t < s - 1. \qquad (A.4)$$

Where s is the signallers actual state.

Below this threshold cost, any signal may function to signal any state. At higher values of H, the inversion of handicaps remains stable, each lower strength class will use a more costly signal as long as

$$H_s - H_{s-1} < \frac{C(1) - C(0) - \frac{1}{2}V}{N} \ \forall s > 1 \quad (A.5)$$

If this condition is met, then the displays will be graded according to cost, with more costly signals indicating lower RHP.