Costly Signaling and Cooperation

Herbert Gintis
Eric Smith
Samuel Bowles*

July 17, 2001

Abstract

We propose an explanation of cooperation among unrelated members of a social group in which cooperation evolves because it constitutes an honest signal of the member's quality as a mate, coalition partner or competitor, and therefore results in advantageous alliances for those signaling in this manner. Our model is framed as a multi-player public goods game that involves no repeated or assortative interactions, so that non-cooperation would be a dominant strategy if there were no signaling benefits. We show that honest signaling of underlying quality by providing a public good to group members can be evolutionarily stable, and can proliferate in a population in which it is initially rare, provided that certain plausible conditions hold, including a link between group-beneficial signaling and underlying qualities of the signaler that would be of benefit to a potential mate or alliance partner. Our model applies to a range of cooperative interactions, including unconditionally sharing individually consumable resources, participating in group raiding or defense, and punishing free-riding or other violations of social norms.

1 Introduction

Cooperation among unrelated individuals has generally been explained by some form of conditional reciprocity (Trivers 1971, Axelrod and Hamilton 1981). However, there is an increasing interest in examining alternative mechanisms for the

---

*Herbert Gintis, Department of Economics, University of Massachusetts, Amherst, hgintis@mediaone.net, http://www-unix.oit.umass.edu/˜gintis. Eric Smith, Department of Anthropology, University of Washington, Box 353100 Seattle, WA 98195-3100, easmith@u.washington.edu, http://faculty.washington.edu/easmith/; Samuel Bowles, Santa Fe Institute and Department of Economics, University of Massachusetts, Amherst, bowles@econs.umass.edu, http://www-unix.oit.umass.edu/˜bowles; We would like to thank Michael Alvard, Carl Bergstrom, Rebecca Bliege Bird, Robert Boyd, Steven Frank, Michael Lachmann, Aimee Plourde, Peter Richerson, Steven Siller, and Richard Sosis for useful comments, and the MacArthur Foundation for financial support. Please do not quote or cite without the authors' permission.

evolution of cooperation (Dugatkin 1997, Pusey and Packer 1997), including mutualism (Brown 1983, Connor 1995), indirect reciprocity (Alexander 1987, Nowak and Sigmund 1998, Wedekind and Milinski 2000, Leimar and Hammerstein 2001), and multilevel selection (Wilson 1977, Bowles 2001, Sober and Wilson 1998, Gintis 2000b). Several authors (Zahavi 1977, 1995, Roberts 1998, Wright 1999, Smith and Bliege Bird 2000,leimar-hammerstein01) have suggested that costly signaling could provide an explanation for cooperation and group-beneficial behavior, but this proposal has not been formally modeled. Following arguments outlined in Smith and Bliege Bird (2000), here we present a game-theoretic model in which cooperative behavior indicates the underlying (dichotomous) quality of the signaler, eliciting a (dichotomous) response from observers that can be mutually beneficial. The model is framed as a multi-player game that involves no repeated or assortative interactions, and assumes a payoff structure that would conform to an multi-player public goods game (and hence universal defection) if there were no signaling benefits. We show that honest signaling of underlying quality by providing a benefit to group members can be evolutionarily stable, and may proliferate when rare as long as high-quality individuals are neither too common nor too rare, and the cost of signaling is sufficiently greater for low than for high quality players.

Our signaling model is distinctive in applying to group (n signalers and n observers) rather than dyadic or many-signaler one-observer interactions. It is also novel in determining endogenously the fraction of the group that signals high quality in equilibrium. Finally, our model operates at the phenotypic level, and abstracts from the specific genetic mechanisms involved in trait-transmission. We supply novel and plausible conditions for the spread of prosocial signals.

We study the case where cooperation involves providing a benefit to all members of the group unconditionally (i.e., without any necessary reciprocation in kind). Given the resulting public goods game payoff structure and the one-shot nature of the interactions, the unique equilibrium of this game under standard assumptions is universal defection, so no player supplies the group benefit. Hence individually costly cooperation could not evolve unless one postulated the group selection of altruistic behavior. Even if interactions among group members were repeated, multi-player cooperation requires implausible forms of coordination, particularly if the group contains more than a few individuals (Boyd and Richerson 1988). The model presented here is meant to apply to such cases, where reciprocity is unlikely to emerge and is vulnerable to free-riding.

We propose that cooperating by providing a benefit to group members may be a reliable signal of signaler quality, where by “quality” we mean genetic or phenotypic attributes that are difficult for others to assess directly, yet have important effects on the payoffs from social interactions with the signaler. Those who provide the benefit to others, or who provide more of the benefit (thus signaling more intensively), honestly advertise their quality as allies, mates, or competitors. This information alters the behavior of other group members to act, for purely selfish motives, in ways that provide positive payoffs to signalers—for example, preferring them as allies or mates, or deferring to them in competitive situations (Smith and Bliege Bird 2000).

The paper is organized as follows. In Section 2 we develop an multi-player model showing that there exists an equilibrium in which quality is honestly signaled over a wide range of parameter values in which low quality types must pay more to signal than high quality types. This model is meant to capture the key features of verbal arguments about cooperation as costly signaling in group-living species, but applies to costly signaling of any kind in multi-player settings. In Section 3 we show that under plausible conditions the costly signaling equilibrium is dynamically stable. In Section 4 we present conditions under which a nonsignaling equilibrium will be displaced by a signaling equilibrium, and in Section 5 we develop a dynamic model determining the movement towards an equilibrium fraction of costly signalers. Given that these results are not specific to cooperation, Section 6 analyzes when costly signaling will take the form of providing benefits to others. A final section draws some conclusions and implications for further research.

2 A Multi-Player Model of Costly Signaling

Consider a group consisting of n members.[1] We think of n as ranging between 10 and 100, to reflect the size of a foraging group or residential band. Once in each period each member of the group can perform an action at personal cost $c > 0$ that confers a benefit $g > 0$ on each other member of the group. We assume that the individual providing g to other group members either does not share in the public good, or else c is the net cost of providing this good, incorporating the providers' share. Since there is a strictly positive cost c to providing the benefit, self-interested individuals will not do so, and those who do provide the benefit will be eliminated by any evolutionary process in which differential replication is monotonic in payoffs. We consider this benefit to be a *public good* in the sense that

$$g(n-1) > c. \tag{1}$$

Our task is to show that providing the benefit may be stable and may evolve. To do this, we add two elements to the above structure of social interactions.

First, suppose group members have a personal characteristic, which we will call 'quality,' that can either be high or low. We assume members know their own quality but not that of others in the group. We assume that the expected cost c per

---

[1] A list of symbols used in this paper is included at the end of the paper.

period of providing the benefit to the group for the high quality type is less than the corresponding expected cost c' for the low quality type, so $0 < c < c'$ (we consider the cases where $c = 0$, $c < 0$, and $c = c'$ later, in Section 2.2).

Second, suppose during a given period each individual has occasion to enter into a profitable alliance (e.g., mating or political coalition) with any one of the other $n - 1$ group members. This other member, whom we will call the Partner, derives a benefit $h = h_o + ag > 0$, where $a \geq 0$, from choosing a high quality ally, and a benefit (or possibly a loss) $l < h$ from choosing a low quality ally, and has payoff zero if no ally is chosen. The parameter a represents the possibility that a Partner may receive greater private benefits from allying with a member who supplies a higher level of the group benefit.

We assume it costs a partner a small amount $\nu > 0$ to monitor signals. For a given Partner, let p be the frequency of high quality members among the group of remaining $n - 1$ members. We assume the Partner knows p but not the high quality/low quality type of individuals among the other $n - 1$ members. A group member receives a payoff $s > 0$ from each of the $n - 1$ Partners who chooses to ally with him, and succeeds in doing so.[2] Finally, we treat each period as a one period game, as would be the case if periods represent generations, or an agent's quality in one period cannot be inferred from his quality in previous periods.

We construct an n-player game in which the players, whom we will call Signalers, choose independently whether or not to signal by providing the benefit to group members. There are four possible Signaler strategies, which we label $\{ss, sn, ns, nn\}$. Here ss means 'always signal, regardless of quality,' sn means 'signal if high quality and do not signal if low quality,' ns means 'do not signal if high quality and signal if low quality,' and finally nn means 'never signal, regardless of quality.' We can abbreviate these as {Always Signal, Signal Truthfully, Signal Untruthfully, Never Signal}.[3]

Each of the n players also plays the role of Partner who must choose an ally (we do not assume that if player i chooses player j as an ally, then player j must choose player i as an ally). A Partner has four possible strategies, which we label $\{aa, ar, ra, rr\}$. Here, using the same convention as with the Signaler, aa means 'Always Accept (whether or not the Signaler signals),' ar means 'Accept if Signaler Signals, Reject if Signaler does not Signal,' ra means 'Reject if Signaler Signals,

---

[2] In an alternative treatment, we would allow the benefit h to a Partner and the benefit s to the group member to be functions of the number of alliances into which the member has entered. It will be clear from the following analysis that our results will continue to hold in this more general setting.

[3] Since in some cases a Signaler may always be high quality or low quality, it may seem unnatural to assign to a Signaler a strategy part of which is never used (e.g., a high quality Signaler does not need an option for the case where he or she is low quality). However Harsanyi (1967) has shown that this formality is harmless. It has the advantage of considerably simplifying the analysis.

Accept if Signaler does not Signal,' rr means 'Always Reject.' We also specify that if no Signaler signals high quality, an ar Partner chooses an ally randomly from the group.[4]

We assume that all Signalers on the one hand, and all Partners on the other, follow the same pure strategy, since it is well known that a mixed strategy equilibrium in an asymmetric game of this type is always unstable (Selten 1980), so it may be ignored.

A distinctive aspect of this model is that individuals signal their quality to all other members of the group, prior to interacting with particular Partners within the group. This assumption reflects our interest in the role signaling might play in favoring collective action (multi-player cooperation) in social species, including humans. But it is often the case that biological signals in other domains such as mate choice, resource competition, and even predator-prey interactions are not private to an intended receiver, but are emitted without the signaler knowing exactly with which among a population of possible observers it might influence. Our model is general enough that it can apply to any domain where costly signaling occurs in an multi-player context.

We represent the probabilistic character of a Signaler's quality by assigning high quality with probability p and low quality with probability $1 - p$. The Signaler, knowing his high or low quality, then chooses either to signal or not, and the other players (in their Partner roles) choose whether or not to consider the Signaler in the pool of potential allies.

It is clear that as long as $ph + (1 - p)l > 0$, there is a nonsignaling equilibrium (nn,aa) in which no one signals and Partners choose allies randomly from all other group members. Similarly, if $ph + (1 - p)l < 0$, there is a nonsignaling equilibrium (nn,rr) in which no agent signals and Partners never choose allies. The *honest signaling equilibrium* is denoted by (sn,ar), indicating that Signalers signal high quality if and only if they are high quality, and Partners choose randomly among those who signaled high quality.

Clearly, it is worthwhile for a Partner to play ar only if the expected benefit from allying with a high quality agent exceeds the cost of monitoring the signal. The payoff from alliances if a Partner chooses randomly from the group is $(hp+l(1-p))$. If we define

$$\delta(x) = 1 - (1 - p)^{x(n-1)}, \tag{2}$$

then, assuming all Signalers are honest (i.e., play ar), the expected payoff to a Partner from monitoring is $h\delta(1) + l(1 - \delta(1)) - \nu$, since with probability $\delta(1)$ the

---

[4] This specification is sensible only if $l \geq 0$, or else a signal-monitoring Partner prefers no ally to a low quality ally. We will assume unless otherwise stated that $l > 0$, thus avoiding a set of parallel derivations that hold when $l < 0$. Our analysis applies equally to both cases.

Partner finds a high quality ally, and with probability $(1-\delta(1))$ the Partner must resort to a low quality ally. Comparing these two payoffs, we see that monitoring is worthwhile for a Partner only if

$$(h-l)(\delta(1)-p) > \nu \tag{3}$$

We assume ν is sufficiently small that this *efficiency of monitoring condition* holds. Note that in most cases $\delta(1) \approx 1$ (e.g., if $p = 0.3$ and $n = 20$, $\delta(1) \approx 0.999$; moreover, $\delta(1)$ is increasing in p and n, so the approximation holds for all larger values of these variables), in which case the efficiency of monitoring condition is approximately $(h-l)(1-p) > \nu$.

2.1 The Honest Signaling Equilibrium

To find the conditions under which there is an honest signaling equilibrium, it is sufficient to derive the conditions under which honest signaling is a best response for one agent, assuming all other members signal honestly (we will presently explore the local stability and global accessibility of such an equilibrium). This gives rise to the matrix shown in Figure 1. Multiplayer games are notoriously unwieldy, so we have introduced several notational simplifications to reduce the clutter of symbols in Figure 1 and the subsequent analysis dependent upon Figure 1.

a. If the fraction of high quality types in the group is p, the fraction of high quality types remaining if the Partner is high quality is $(pn-1)/(n-1)$ and is $pn/(n-1)$ if the Partner is low quality. We assume n is sufficiently large that both these numbers can be approximated by p.

b. We have multiplied the payoff to Signalers from alliances by n to reflect the fact that a Signaler pays the cost c only once per period, but has n potential Partners (actually $n-1$ but again we assume this is n to reduce notational overhead), each of whom can independently ally with the Signaler.

c. We have not included the benefit to others that accrues to both players from the actions of the other $n-2$ players, since this payoff merely adds a constant to each row of the game matrix for the Partner, and a constant to each column of the matrix for the Signaler. To see this, suppose k of the $n-2$ remaining players provide the benefit to group members. Then both Signaler and Partner receive payoff kg from the signalers' provision of benefits, no matter what strategies they choose. Moreover, whatever strategy the Signaler chooses, the Partner receives the same benefit from the Signaler. For instance, if the Signaler chooses sn, then the Partner receives a payoff of pg from the Signaler's provision of benefits, whether Partner chooses aa, ar, ra, or rr. Since adding a constant to the payoffs of a player, given the choices of the other players, cannot change the player's best response strategy, we omit these payoffs.

d. We assume that n is sufficiently large and p is sufficiently far from zero, that the probability that there is at least one high quality player is unity (i.e., we assume $\delta(1) = 1$).

	aa	ar	ra	rr
ss	$s - pc - (1-p)c'$ $ph + (1-p)l$	$s/p - pc - (1-p)c'$ $ph + (1-p)l - v$	$-pc - (1-p)c'$ $-v$	$-pc - (1-p)c'$ 0
sn	$s - pc$ $ph + (1-p)l$	$s - pc$ $h - v$	$s - pc$ $l - v$	$-pc$ 0
ns	$s - (1-p)c'$ $ph + (1-p)l$	$(1-p)(s/p - c')$ $l - v$	$s - (1-p)c'$ $h - v$	$-(1-p)c'$ 0
nn	s $ph + (1-p)l$	0 $-v$	s $ph + (1-p)l - v$	0 0

Figure 1: The Matrix of Payoffs to a Signaler S and a Partner P, assuming all other Signalers play the same strategy as S.

To illustrate how the entries in Figure 1 are calculated, we will derive them for the honest signaling equilibrium, which is the highlighted (sn,ar) box. In this case, a high quality Signaler provides the benefit at cost c and receives expected benefit s/pn from each Partner, since each Partner now chooses randomly from among the group of pn Signalers who provided the benefit. Multiplying the alliance payoff by n, because there are n Partners, and multiplying the net payoff by p, since the Signaler is high quality with probability p, the expected payoff to the Signaler is $p(s/p-c) = s - pc$. Similarly, since a Signaler can have multiple allies, a Partner certainly finds an ally among the high quality members, giving payoff $h - v$.[5] A similar argument is used to fill in the other entries in Figure 1.

From the construction of the matrix in Figure 1, we know that a pair of best responses for Signaler and Partner determine a Nash equilibrium of the game, since if honest signaling is a best response for one Signaler when all other signalers use honest signaling, then the same is true for any signaler. It follows that the conditions

---

[5] The assumption that a Partner can ally with only one signaler is somewhat arbitrary, but it is completely straightforward to extend our analysis to the case where Partners can ally with multiple Signalers. It is somewhat more challenging to allow the Partner's alliance benefit h to depend on the number of allies possessed by the Signaler with whom he allies. On grounds of simplicity, we will avoid this more general treatment.

$pc' > s > pc$, and $h > l$ are necessary and sufficient for honest signaling (sn,ar) to be a strict Nash equilibrium. The conditions for this signaling equilibrium to exist are easily interpreted, and minimally what one would *prima facie* expect in a costly signaling framework. The first, $s > pc$, requires that the benefits of signaling exceed the expected cost of signaling for the high quality type. The second, $s < pc'$, requires exactly the opposite for low quality types. Finally, $h > l$ states that the benefit from allying with a high quality type is greater than with a low quality type.

2.2 Frequency Dependence of the Honest Signaling Equilibrium

We have the following theorem.

Theorem 1. *Suppose there is a positive payoff to an alliance for the Signaler ($s > 0$), and Partners prefer to ally with high quality types ($h > l$). Then there is a range of frequencies p of high quality types for which there is an honest signaling equilibrium if and only signaling is more costly for low quality than high quality types ($c' > c$), and more costly than the payoff to a single alliance ($c' > s$).*

To prove the theorem, we rewrite the condition $pc' > s > pc$ as

$$c' > \frac{s}{p} > c. \tag{4}$$

If (4) holds, then clearly $c' > c$ and since $p < 1$, we must have $s < c'$. Conversely, if these parameter inequalities hold, then (4) holds for any $p > 0$ in the interval $(s/c', s/c)$. The efficient monitoring condition (3) is sufficient to ensure that ar is a best response for the Partner.

Note that this theorem does not require $c > 0$. If $c = 0$, so signaling is not costly for the high quality signaler, there still exists an honest signaling equilibrium for some p, so long as $c' > s$. Indeed, it is easy to see that we can have $c < 0$ and the signaling equilibrium will still exist under the same conditions. This is an important observation because it implies that the cost of signaling for the high quality type cannot be signed, so even signals that are intrinsically beneficial to the Signaler can be part of an honest signaling equilibrium.

The reader will note that the equilibrium conditions for honest signaling that we have just derived are *dependent upon the frequency p of high quality types in the group*. This aspect of costly signaling has of course been noted in verbal descriptions of costly signaling, but did not appear in the Grafen's pioneering contribution (Grafen, 1990a,b) nor in most of the more recent papers on the topic (but see Lachmann and Bergstrom (1998), and Siller, 1998).

2.3 Honest Signaling by Providing a Public Good Enhances Average Fitness

To see that honest signaling by providing a public good is fitness enhancing for the group as a whole, suppose that every member of the population is both a Partner and a Signaler in each period. Then average payoffs for the two stable equilibria are just the sum of the entries in the reduced normal form matrix for the relevant strategy profile, plus the gains to all parties from the provision of Signaler benefits. Thus the honest signaling equilibrium has higher average payoffs if $h+(n-1)pg+p(s-c) > hp+(1-p)l$, which reduces to

$$p\left(1-\frac{(n-1)g+s-c}{h-l}\right) < 1, \tag{5}$$

which, by (1) is true if the benefit g is a public good. Note that for the honest signaling equilibrium to have higher average payoffs, g need not represent a public good. Even conferring harm ($g < 0$) will entail higher average payoffs if a sufficiently large number of alliances with high quality individuals result.

2.4 Cost, Benefits, and Punishment

Our model assumes low quality types have higher signaling costs. An alternative assumption that is sometimes more accurate (Godfray 1991, Maynard Smith 1991, Johnstone 1997, Getty 1998) is that both types face the same marginal signaling costs, but high quality types reap higher marginal benefits per unit of signaling than do low quality types. In terms of our parameters, this means $c = c'$ and $s > s'$, where s' is the value of an alliance to a low quality signaler. If we solve for Nash equilibria using the parameters c, c', s, and s' with $c, c' > 0$, we find that the conditions for an honest signaling equilibrium are simply $s > pc$, and $s' < pc'$. These conditions of course reduce to the above conditions when $s = s'$, and we will not pursue this variant of the model further in this paper.

Among the forms of signaling described by the model is the punishment of those who violate group-beneficial norms. We offer the following brief account to illustrate how our model captures this form of signaling. Suppose that a group of n members can cooperate to provide some group benefit. By cooperating, each member contributes a total benefit of b to others at a fitness cost of c to himself. Thus, the gain from defecting is $c-b/n$ and to induce cooperation, members must be punished at least $c-b/n$ for defecting. Now suppose that a high quality individual can impose $c-b/n$ on defectors at a personal cost of c, whereas a low quality individual must incur cost $c' > c$ to achieve the same effect. Following the model presented above, there will be an equilibrium in which high quality individuals will punish and low quality ones will not, provided certain parameter values obtain. We will discuss the significance of the punishment form of costly signaling in Section 7.

3 Dynamics

We make this model dynamic by assuming that the fraction of honest Signalers increases when the payoff to honest signaling exceeds that of Never Signal, and the fraction of ar Partners increases when the payoff to ar exceeds that of aa. We model these as a‘replicator dynamic’ (Taylor and Jonker 1978), which means that the rate of growth of a Signaler strategy, represented as a fraction of Signalers in the population using that strategy, equals the difference between the payoff to that strategy and the average payoff to all Signaler strategies, and similarly for Partner strategies. This dynamic can reflect either cultural change, in which members are prone to switch from inferior to superior strategies, or genetic change, in which those who pursue successful strategies have more offspring, who tend to follow their parent’s strategies.[6]

In analyzing the dynamics of our system, since we are now considering the whole range of values of $\alpha \in [0, 1]$, we can no longer abstract from the probability of finding an ally schedule $\delta(x)$. We must calculate the expected payoffs when the fraction of sn Signalers is α and the fraction of ar Partners is β. We derive these payoffs as follows.

With probability p an honest signaler indicates that he is high quality. The number of Partners of type ar is $\beta(n-1)$. Since the expected number of signalers who signaled high quality is $\alpha p(n-1)$, the expected payoff for each such signaler is $\beta(n-1)s/\alpha p(n-1)$. We multiply this by p, the probability that an honest signaler is of high quality to arrive at the first term for π_{sn} below. With probability $1-p$ an honest signaler is not of high quality and does not signal. This signaler may be chosen as an ally of an ar Partners when there are zero high quality signalers. There are $\beta(n-1)$ Partners of type ar, and the probability there are no high signalers is $(1-\delta(\alpha))$, so the expected number of alliances of this type is $\beta(n-1)(1-\delta(\alpha))$, which must be shared by the $n-1$ signalers. This accounts for the second term in π_{sn} The expected number of alliances with aa types is $(1-\beta)(n-1)$, shared among the $n-1$ signalers, giving the third term in π_{sn}. The final term in the expression is the expected cost of signaling high quality.

Similarly, the first term for the expected payoff to Never Signal (π_{nn} in the equations below) is s times the probability that a non-signaler allies with an ar-type when no agent signaled high quality, and the second term is s times the probability that a signaler makes an alliance with an aa-type Partner. The payoff to Always

---

[6] For various derivations of the replicator dynamic equations, see Gintis (2000a), Chapter 9.

Accept is the expected payoff to an alliance, $ph + (1 - p)l$. Finally, the expected payoff to an ar-type is h times $\delta(\alpha)$, which is the probability of making an alliance with someone who signaled high quality, plus the probability of an alliance with someone who did not signal high quality, $1 - \delta(\alpha)$, times the average quality of such an agent, who is of average quality, $hp+(1-p)l$ with probability $(1-\alpha)$, and is a honest signaler of low quality with probability α. We thus have the following expected payoffs:

Signal Honestly:

$$\pi_{sn} = \frac{\beta s}{\alpha} + (1 - p)\beta s(1 - \delta(\alpha)) + (1 - \beta)s - pc$$

Never Signal:

$$\pi_{nn} = \beta s(1 - \delta(\alpha)) + (1 - \beta)s$$

Always Accept:

$$\pi_{aa} = ph + (1 - p)l$$

Accept if Signaler Signals:

$$\pi_{ar} = \delta(\alpha)h + (1 - \delta(\alpha))(\alpha l + (1 - \alpha)(ph + ql)) - \nu$$

Average Signaler Payoff:

$$\overline{\pi}^1 = \alpha\pi_{sn} + (1 - \alpha)\pi_{nn}$$

Average Partner Payoff:

$$\overline{\pi}^2 = \beta\pi_{ar} + (1 - \beta)\pi_{aa}$$

The replicator equations are then

$$\dot{\alpha} = \alpha(\pi_{sn} - \overline{\pi}^1) \tag{6}$$

$$\dot{\beta} = \beta(\pi_{ar} - \overline{\pi}^2), \tag{7}$$

which reduce to

$$\dot{\alpha} = \alpha(1 - \alpha)(s\beta(1 - \alpha(1 - \delta(\alpha)))/\alpha - cp)$$

$$\dot{\beta} = \beta(1 - \beta)((h - l)(\delta(\alpha) - p + (1 - \alpha)p(1 - \delta(\alpha))) - \nu).$$

These equations express the familiar result that the rate of change of the frequency of a trait in a population varies with the variance of fitness, or equivalently, the variance of the trait times the effect of the trait on fitness. The first says that the rate of increase of honest signaling equals the variance of the frequency of honest signaling, which is $\alpha(1 - \alpha)$ times the net gain from honest signaling. The second says that the rate of increase in ar equals the variance of the frequency of ar, which is $\beta(1 - \beta)$, times the net gain from ar.

The replicator equations (6) and (7) have five equilibria. Four correspond to $\alpha = 0, 1$ and $\beta = 0, 1$, and the fifth is a mixed strategy equilibrium in which all four types of agents are present. The Jacobian matrix of the replicator equations is a quite complicated expression, but its eigenvalues for $\alpha = \beta = 0$ are $\{-cp, -v\}$, which are strictly negative assuming the cost of signaling, c, is positive. Thus the no-signaling equilibrium is stable under this condition. At the equilibrium $\alpha = 0$, $\beta = 1$ the Jacobian has eigenvalues

$$\{-cp - ps, v\}, \tag{8}$$

indicating instability. At the equilibrium $\alpha = 1, \beta = 0$ the Jacobian has eigenvalues

$$\{cp, (h - l)(\delta(1) - p) - v\}. \tag{9}$$

Thus this equilibrium is also unstable for $c > 0$, or when the efficient monitoring condition holds. At the truthful signaling equilibrium, corresponding to $\alpha = \beta = 1$, the Jacobian has eigenvalues

$$\{cp - (1 - (1 - \delta(1))p)s, v - (h - l)(\delta(1) - p)\}, \tag{10}$$

both of which must be strictly negative for honest signaling to be a stable equilibrium. The second expression is negative by the efficient monitoring condition 3. The first expression is negative when

$$\frac{c}{s} < (1 - (1 - \delta(1))p), \tag{11}$$

which becomes the criterion $s > cp$ for an honest signaling equilibrium, given the approximation $\delta(1) \approx 1$. We have seen that this is the case so long as p is not too small and n is sufficiently large. Assuming this, both eigenvalues in (10) are negative, and the honest signaling equilibrium is stable. We thus have

Theorem 2. *Suppose the conditions for an honest signaling equilibrium hold (see Theorem 1) with frequency p of high quality types. Then if (11) and (3) hold and $c > 0$, both the nonsignaling and the honest signaling equilibria are stable. Moreover, there are three unstable equilibria and a ridge line connecting them that separates the basins of attraction of the two stable equilibria.*

Figure 2 shows the phase diagram for the dynamical system. The equilibrium indicated by B in Figure 2 is unstable, as can be confirmed by the vector field derived from the replicator equations (6) and (7) indicated by the arrows.

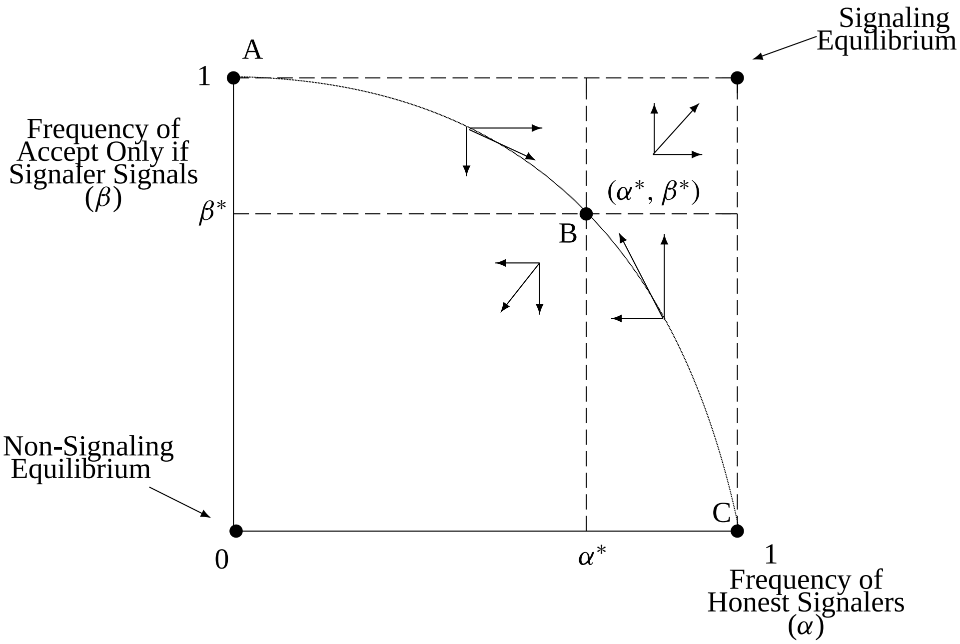

Figure 2: The Phase Diagram. Points below the ridge line ABC are in the basin of attraction of the Nonsignaling equilibrium.

4 The Evolution of Signaling

The analysis thus far has revealed that while an honest signaling equilibrium will exist under a broad range of parameter values, so will a non-signaling equilibrium. Are there plausible mechanisms leading from the nonsignaling to the signaling equilibrium? We can offer three complementary mechanisms of this type.

4.1 Stochastic Shocks

A population at the nonsignaling equilibrium might be displaced by a series of stochastic events into the basin of attraction of the signaling equilibrium, should the underlying parameters be such that the honest signaling equilibrium exists. Since a small number of Signalers and signal-monitoring Partners can invade a nonsignaling group, the movement from a nonsignaling to a signaling equilibrium is much more likely than the reverse movement, which requires the simultaneous extinguishing of signaling and signal monitoring of most group members. Having thus proliferated within a single group or a few groups, signaling equilibria may proliferate in a larger population through multilevel selection.

4.2 Mutualism

Throughout this paper we have assumed signaling high quality is costly to the high quality agent; i.e., $c > 0$. But as Theorem 1 shows, the existence of an honest signaling equilibrium does not depend on this fact. The condition $c < 0$ indicates that providing the benefit g is beneficial to the high quality Signaler, and hence g is a form of *byproduct mutualism*. Studies of animal behavior indicate the importance of byproduct mutualism (Dugatkin 1996, Milinski 1996, Dugatkin and Mesterton-Gibbons 1996, Mesterton-Gibbons and Dugatkin 1997), which doubtless plays a role in human societies as well.

When $c < 0$, an inspection of the replicator equations (6) shows that only the honest signaling equilibrium ($\alpha = \beta = 1$) is stable. Indeed, in this case a mutant honest signaler has higher fitness than a nonsignaler, even in the absence of Partners who monitor and respond to the signal. Selection will then favor Partner strategy ar, and the honest signaling equilibrium will obtain in the long run.

Once a signaling equilibrium is attained, provided the conditions for prosociality developed below in Section 6 obtain, there will be a tendency for signals that provided high social benefits to displace signals providing low social benefits, even when the costs associated with these benefits are strictly positive. Thus, even if the highly prosocial signals involve positive signaling costs, such signals will evolve in the long run.

4.3 Inclusive Fitness

Another mechanism leading from nonsignaling to signaling equilibria involves the generalization of behavior from highly genetically related groups to groups of unrelated individuals.

Suppose our group of n members has degree of relatedness $r > 0$. Then the inclusive fitness cost of honest signaling for a Signaler is $c = c_o - (n-1)gr$. If signaling is prosocial and c_o is not too large, we will have $c < 0$, so the analysis of the previous section applies: positive relatedness facilitates the evolution of prosocial honest signaling.

Once the honest signaling equilibrium is established in a group of related individuals, migration in and out of the group will not destroy the honest signaling equilibrium, provided p remains in the appropriate range and the other conditions specified in Theorem 1 obtain. Hence, the honest signaling equilibrium will persist even under conditions of zero relatedness.

5 Heritability of Quality

We have shown that signaling by providing benefits to members could proliferate when rare, and would under plausible conditions be sustained in a population in which behaviors evolved according to the payoff-monotonic updating described in equations (6), (7). But this is not sufficient to ensure the evolutionary success of such signaling. Recall that the existence of the signaling equilibrium requires that the high quality types not be excessively prevalent in the population. But if high quality types have higher fitness than low quality types in the signaling equilibrium, and if quality is heritable, their frequency p may increase over time, thus undermining the signaling equilibrium.

The payoff difference between high and low quality types in the signaling equilibrium is $s/p - c > 0$ so in the absence of any other influence on p, high quality would evolve to fixation. But we have modeled only a subset of the influences on p, and we may suppose other influences to be at work. The relevant differential equation is

$$\dot{p} = p(1-p)(s/p-c) - zp + w(1-p) \tag{12}$$

where the first term expresses the rate at which the differential fitness of high quality types is translated into offspring, $z > 0$ is the proportion of offspring produced by high quality parents that are low quality, and $w > 0$ is the proportion of offspring produced by low quality parents that are high quality. Equation (12) can be simplified to

$$\dot{p} = cp^2 - (c+s+w+z)p + s + w. \tag{13}$$

Since the left hand side of (13) is $s + w > 0$ at $p = 0$ and $-z < 0$ at $p = 1$, there is surely a stable equilibrium p^*, where $0 < p^* < 1$. We have

$$p^* = \frac{c+s+w+z-\sqrt{(c+s+w+z)^2 - 4c(s+w)}}{2c}.$$

This expression is complicated, but it does have some intuitive implications. Setting the right hand side of (12) to zero and totally differentiating to see how p^* varies in response to changes in our parameters, we find that (a) an increase in the cost c to the high quality type of signaling leads to a lower equilibrium value of p^*; (b) an increase in the benefit s of an alliance to the Signaler raises the equilibrium value of p^*; (c) an increase in rate w at which low quality agents produce high quality offspring raises the equilibrium value of p^*; and an increase in rate z at which high quality agents produce low quality offspring lowers the equilibrium value of p^*.

There are of course various social and biological mechanisms that could lead to positive levels for either z or w. For example, the characteristics that confer high quality or low quality could be purely or substantially environmentally determined, or determined by cultural factors that are passed on through biased or horizontal rather than vertical cultural transmission (Cavalli-Sforza and Feldman 1981, Boyd and Richerson 1985). Or, the fitness benefits associated with prosocial signaling of high quality may be confined to some period of the life cycle, and be partially offset by disadvantageous effects at other ages.

We have explored the global dynamics of our population, showing first, that honest signaling may proliferate when rare, and second that an honest signaling equilibrium once attained need not set in motion a rise in the fraction of high quality types that would eventually destroy the conditions for an honest signaling equilibrium. It remains to be shown that in an honest signaling equilibrium group beneficial signals would be favored over other honest signals of quality.

6 Why Signal by Providing Benefits to Others?

We have shown that when signaling takes place in an n-player social context, and providing benefits to other serves as an honest signal of underlying quality, the payoffs to signaling can solve the problem of maintaining unilateral cooperation in a group when self-interested agents would not otherwise provide the benefits. In other words, under conditions standard to the costly signaling framework and specified in Section 2, if the social benefits that signalers receive from signaling exceed the costs of signaling when they are high quality ($s > pc$), these can provide a net advantage for what would otherwise be altruistic contributions to the group. However, such group-beneficial signaling is only a special case of the model presented in Section 2. Indeed, in our model, the existence of a costly signaling equilibrium does not require that there be any group-beneficial effects of the signal. To see this, note that the per member benefit g could equally well be positive, negative, or zero without altering the conditions for existence of stability of a costly signaling equilibrium. Moreover, if $ph + (1 - p)l > 0$, whenever there is a stable costly signaling equilibrium, there is another stable equilibrium in which that signal is not used. Therefore, when there are a variety of possible signals, our analysis to this point does not tell us which among them will in fact be favored.

Let us call a costly signal *feasible* if it satisfies the conditions for a costly signaling equilibrium developed in Section 2. We have shown that prosocial signals can meet these conditions, but we have not shown that there is any particular reason to expect feasible costly signals deployed within a group to be prosocial. We can specify several conditions under which prosocial signaling will be favored over alternative signals.

6.1 Signaling Prosocial Quality

First, suppose there are several feasible signals and consider an honest signaling equilibrium in which the most prosocial signal (i.e., the signal that provides the highest level of benefit, g^*, to group members) is not used. Then as long as $a > 0$, the benefit $h^* = h_o + ag^*$ to a mutant Partner who allies with a mutant Signaler who provides the public benefit at level g^* will be higher than that of Partners who recognize only signals with public good value $g < g^*$. This being the case, the Partner strategy of allying with Signalers providing g^* will increase in frequency according to a straightforward adaptive dynamic.[7] This in turn increases the payoff to g^* signalers, since they now achieve more alliances. Thus, as long as the marginal gain to the signaler from increased alliances is greater than the added cost of signaling by providing g^* rather than g, the equilibrium involving the less prosocial signal is not stable against invasion by more prosocial Signalers and Partners who respond to their signals.

The assumption that $a > 0$, on which this argument hinges, asserts that the level of public benefit provided is positively correlated with the expected level of benefit the Signaler provides to the Partner. This argument is analogous to the "direct benefits," or "good parent," explanation for female preference of males who signal superior ability to provide parental care or other resources (Johnstone 1995, Iwasa and Pomiankowski 1999). In both cases, a high-quality individual is more likely to provide the social benefit because the cost of doing so is lower than for a low-quality individual, and the individual who provides the most prosocial benefit will also be equipped to supply the most private benefit to the observer of the signal. In this case, the group benefits produced by the Signaler are incidental to the Partner's preference, but because they produce direct benefits to the Partner they are valued in of themselves, not simply as indicator traits.

One example of a positive correlation between the social value of a signal and its value to a potential ally is the case where a high value of g reflects the Signaler's willingness to act in a generous manner in bargaining over the distribution of the benefits of his efforts. This behavioral propensity will likely generalize to the signaler's private alliances. Alternatively, a high value of g might reflect the signaler's capacity to honor commitments by demonstrating the ability to share benefits that would not be worth providing if the Signaler did not intend to continue the alliance in the future (Schelling 1978, Zahavi and Zahavi 1997).

Another example of a signal that is both prosocial (providing group benefits) and directly beneficial to allies is signaling extraordinary foraging ability through

---

[7] By an adaptive dynamic we mean a situation in which, if a parameter positively affects fitness, over time strategies exhibiting higher values of this parameter will displace strategies exhibiting lower values (Dieckmann 1997, Hofbauer and Sigmund 1998).

unconditional sharing of surplus resources. This will provide direct benefits to Partners if Signalers also use this ability to provision their allies (e.g., mates with dependent offspring). A final example is defending the group against enemy attack. This signals qualities such as strength and fighting ability, which will provide direct benefits to Partners if Signalers with such attributes are more likely to defend allies in internal disputes. Again, we are assuming that a high-quality individual signals by providing these collective and individual benefits because the cost of doing so, c, is lower than it would be for a low-quality individual, c'.

It is of course possible to construct scenarios where high-quality individuals signal in socially neutral or even antisocial ways. For example, one could signal extraordinary foraging ability by focusing on resources that require great skill to capture but provide little food value (e.g., the spearfishing pursued by some Meriam men, as described in Bliege Bird, Smith and Bird 2001), or by conspicuously consuming surplus resources (Veblen 1899), or even by destroying them (as in some competitive potlatches among the Kwakiutl Indians, Ruyle 1973). Similarly, strength and fighting ability could be signaled by bullying members of one's group, or engaging in repeated brawls. We do not discount these possibilities, but simply argue that when prosocial acts are effective at signaling underlying qualities, and also provide direct benefits to potential allies (enhancing h), Partners will prefer to ally with Signalers using such signals, thus preferentially benefiting them (*via* s).

6.2 Reciprocity

Suppose that $a = 0$, so the above argument does not hold, but suppose all else equal, Partners prefer to ally with the agents who have conferred the highest benefits upon them. Notice that this behavior imposes no cost on the mutant Partner who responds to a more prosocial signal, so there is no fitness penalty associated with mutant Partners who show a preferences for Signalers who have provided them with greater public benefits. In this case once again the equilibrium with group benefit g can be invaded by a mutant signaler that uses the most prosocial signal g^*. Mutant signalers of this type will thus increase in frequency under the operation of an adaptive dynamic. This will continue until the most prosocial feasible signal displaces the antisocial or less prosocial signal.

6.3 Broadcast Efficiency

Another mechanism favoring prosocial signaling is that signaling by providing a benefit to group members may increase "broadcast efficiency" (Bliege Bird 1999, Smith and Bliege Bird 2000), in the sense that it attracts a larger audience to witness the signal. The broadcast efficiency argument could be formalized in a variety of ways. One simple approach that captures the essential point is as follows. We initially assumed that all members of a Signaler's group "see" the signal. But suppose there are several signals $\{\sigma_1, \ldots, \sigma_k\}$ each of which is observed by a fraction $\{f_1, \ldots, f_k\}$ of the group. We call f_i the "broadcast strength" of signal σ_i. Then the payoffs to the Signaler using σ_i will be reduced by $(1 - f_i)s$ in a signaling equilibrium, because the Signaler now receives an expected benefit only from the $f_i(n - 1)$ members who "viewed" the signal. Therefore, everything else being equal, Signaler strategies that maximize the fraction of the group that views the signal will gain higher payoffs and will increase their share in the population. If a prosocial signal attracts a larger audience, the higher broadcast strength per unit signal cost (f_i/c) will favor signaler strategies of this type. Detailed study will be needed to determine the degree to which this applies to particular cases.

The broadcast efficiency argument applies when prosocial signals attract a larger audience than alternative signals of equivalent cost. Signals that consist of providing a consumable public good (e.g., foods of sufficient value or rarity to attract audiences) will plausibly have this effect, while prosocial acts such as punishing non-cooperators may not. An example of the former case involves chimpanzees (*Pan troglodytes*), who live in social groups that travel and forage in a "fission-fusion" pattern (Wrangham 1980), so that potential signal observers are often within close range but not necessarily in visual contact. While most foraging is non-cooperative, and most consumption is by the acquirer, an important exception is hunting (Stanford 1999). It is interesting to note that among wild chimpanzees, extra-kin food sharing only involves group-hunted resources, and the hunters are always males. Stanford, Wallis, Mpongo and Goodall (1994) found that the best predictor of hunting frequency is the number of estrous females present in the social group, even though females are not the primary recipients of meat shares, whereas Mitani and Watts (2001), working with a different population, found that in a multivariate analysis only the number of adult males was a good predictor. Stanford interprets his results as supporting meat-for-sex exchange, while Mitani and Watts favor an explanation involving meat-for-male coalition support exchange. However, the available evidence is fully consistent with a signaling/broadcast efficiency argument, which does not require the direct trade of Stanford's scenario and avoids the problems of enforcing reciprocity that arise in Mitani and Watt's.

Costly signaling has been proposed as an explanation for certain types of food-sharing in human societies, such as providing large and/or difficult-to-harvest game, or large quantities of food for consumption at ritual feasts (Boone 1998, Gurven, Allen-Arave, Hill and Hurtado 2000, Hawkes, O'Connell and Blurton Jones 2001, Smith and Bliege Bird 2000, Sosis 2000). In most such cases, as in the chimpanzee case, there is not sufficient information available to judge if key conditions for costly signaling, such as quality-dependent signaling, are present. In the case studied by Bliege Bird and colleagues (Bliege Bird, Smith and Bird 2001), such data are now available, and agree with costly signaling predictions. This study also shows that the most prosocial form of signaling—unconditionally sharing large game—has high broadcast efficiency, and is more likely to be done through feasts than through household-to-household sharing. It is clear that young men are more likely to establish reputations for foraging ability by providing large game (marine turtles) for feasts attended by upwards of 200 people (mean $= 174.9$, $n = 54$ feasts) than by other means, such as conspicuous consumption or minimizing foraging time to supply domestic needs, which would only be observed by immediate neighbors, and even for them be less conspicuous than feast contributions.

6.4 Group Selection of Alternative Equilibria

A plausible process favoring the selection of prosocial costly signaling, though not modeled in this paper, is a process of groups selection among alternative local equilibria modeled by Boyd and Richerson (1990) and others. The model presented in Section 2 has both an honest signaling and a nonsignaling equilibrium, both stable in the replicator dynamic. Which equilibrium would occur in a particular case will depend on local circumstances, including social and environmental factors determining the level of key parameters, as discussed in Section 3, and perhaps on initial composition of the population. Given this variation in the metapopulation, it follows that local groups will vary in the level of prosocial costly signaling, and that groups with a high level will have members who, on average, are more fit than groups in which such behavior is absent. Such groups, by withstanding extinction and dispersion, and by having superior strength in hostile interactions with other groups, can spread the prosocial practices beyond their original boundaries (Gintis 2000b). Coupled to such a framework, our model provides a possible basis for a more general understanding of which among the multiplicity of signals are likely to evolve and persist, namely, those characterized by a large basin of attraction for the associated equilibria.

6.5 Summary of Forces Favoring Prosocial Signaling

We have offered several distinct and complementary reasons why prosocial signaling may be favored over other forms of signaling. First, Partners may receive higher private benefits from signalers who use more prosocial signals. Second, potential Partners may simply prefer agents who have conferred the highest group benefit g upon them—a preference with no individual level fitness-reducing consequences and strong group-level fitness increasing consequences. Third, some prosocial signals may attract larger audiences than other signals of equal cost, and hence have higher broadcast strength. Fourth, group selection among alternative equilibria can favor outcomes with higher group fitness, which can give an evolutionary advantage to equilibria with more prosocial effects.

7 Conclusion

Costly signaling of quality for purposes of mating, alliance formation, and warning off potential enemies has been reported in many species, including humans (Zahavi 1977a, Grafen 1990a, Maynard Smith 1991, Johnstone 1995, Wright 1999). We have proposed a multiplayer game-theoretic model of costly signaling and shown that under plausible parameter values, a class of signals that themselves contribute to group benefits may proliferate in a population when rare, and constitute evolutionary stable strategies. Costly signaling may thus provide a mechanism for the evolution of cooperative and other group beneficial practices capable of working independently of repeated interactions, positive assortment, and multilevel selection, though these latter factors may act to reinforce such evolution.

While existing game-theoretic models of signaling are framed as dyadic or many-signaler one-observer interactions (e.g., Johnstone, 1999), our signaling model is distinctive in applying to group interactions. It is also novel in determining endogenously the fraction of the group that signals high quality in equilibrium. We show that honest signaling of underlying quality by providing a benefit to group members can be evolutionarily stable, and may proliferate when rare as long as high-quality individuals are neither too common nor too rare, and the cost of signaling is sufficiently greater for low-quality than for high-quality players.

Our model is general enough to apply to a range of social interactions. First, the benefit whose provision signals high quality may take the form of individually consumable resources. For example, the widespread practice among hunter-gatherers of sharing individually harvested resources unconditionally with all members of the community has presented a puzzle for models based on reciprocity and risk reduction (Hawkes 1993). Some hunters consistently provide more than others while sharing equally in the catch. These “altruistic” providers in fact reap higher social status and reproductive success than their less-productive peers, despite the absence of any conditional exchange of “meat for mates” (Kaplan and Hill 1985, Marlowe 1999, Bliege Bird et al. 2001). Our model formalizes the conditions under which a costly signaling explanation might account for such a pattern. Catching large game, the most commonly shared resource, requires skill and endurance, and readily attracts an audience to consume it, allowing the signals sent by hunting success and generosity to be broadcast efficiently. Whether hunting success is a condition-dependent signal of quality is not yet firmly established, though some current work matches this prediction (Bliege Bird et al. 2001).

The model also can apply to the provisioning of public goods that are not individually consumable. For instance, participating in group raiding or defense—an individually costly behavior that is common among chimpanzees as well as human societies (Boehm 1992)—provides benefits that are available to all group members and cannot be hoarded. Similarly, punishment of those who free-ride or otherwise engage in anti-social behavior is a critical mechanism for enforcing cooperation in many social contexts (Frank 1995). The role that costly signaling might play in enforcement of prosocial behavior is as yet untested, but deserves further investigation.

Our model applies as well to situations involving punishing those who free ride on the group's cooperative activities. It is well known that, while enforcing cooperation by punishing defectors can solve collective action problems, such enforcement itself provides a personally costly public good, and thus poses a second-order collective action problem (Hardin 1982). Boyd and Richerson (1992) demonstrated that if enforcement takes the form of punishing both non-cooperators and non-punishers, then cooperation (or anything else) can be evolutionarily stable, even in large groups. Such punishment may be a potent element in stabilizing cooperation in many types of social systems (Boyd and Richerson 1992, Clutton-Brock and Parker 1995, Frank 1995, Michod 1997, Fehr and Gächter 2000). The model presented in Section 2 provides one mechanism for the evolution of such a system. In this version, enforcement—punishment of non-cooperators—itself is the benefit to others that signals high quality. Our model easily allows such punishment or enforcement to serve as the costly signal, and hence to be maintained when the conditions for evolutionary stability specified in the model are met, as discussed in Section 2.4.

Honest signaling of quality need not be group beneficial, of course, and our signaling model applies equally well to socially neutral or harmful forms of costly signaling, such as conspicuous private consumption, brawling and dueling, flouting social norms with impunity, and the like. We suggested several factors that might make signals with prosocial consequences (such as public generosity) more likely to evolve than equally condition-dependent neutral or harmful signals: the greater likelihood that observers will ally with the signaler (due to present and expected future direct benefits), the likelihood that prosocial signals that take the form of individually consumable public goods will attract a larger audience for the signaler, and the operation of group selection on multiple alternative equilibria. We related the first two of these to parameters of our model. However, we do not claim on the basis of this model (nor any existing theory) that prosocial signals will always have an evolutionary advantage over signals without group benefits. Indeed, such a claim would fly in the face of the ample empirical evidence for individually adaptive but collectively harmful displays among humans and many other species. Rather, we suggest that several factors (just summarized) may make the evolution of prosocial signals more likely, if the relevant conditions pertain in particular cases. Our model specifies the conditions that can produce an honest signaling equilibrium in multi-player social contexts, and illuminates the additional factors that could tip the balance toward or away from prosocial signals per se. Further theoretical and empirical work is warranted to determine the fruitfulness of this approach to the evolution of cooperation.

List of Symbols

This is a list of symbols used in places other than where they first defined.

a = the weight of the Signaler benefit in a Partner's alliance payoffs: $h = h_{\text{o}} + ag$, $l = l_{\text{o}} + ag$
c = signaling cost for a high-quality type
c' = signaling cost for a low-quality type
f = probability a signal is perceived (broadcast strength)
g = individual benefit conferred on each group member by Signaler
h = Partner's payoff from allying with a high quality type
l = Partner's payoff from allying with a low quality type
n = group size
p = fraction of population who are high-quality types
s = Signaler's payoff from allying with a Partner
w = fraction of offspring of low-quality parents who are high-quality
z = fraction of offspring of high-quality parents who are low-quality
α = fraction of Signalers who signal honestly
β = fraction of Partners who prefer to ally with Signaler who signals
δ = probability that a Partner will successfully ally
ν = fitness cost to Partner from monitoring signal
π = expected payoff for a given (subscripted) strategy
σ_i = signal i

REFERENCES

Alexander, R. D., *The Biology of Moral Systems* (New York: Aldine, 1987).

Axelrod, Robert and William D. Hamilton, "The Evolution of Cooperation," *Science* 211 (1981):1390–1396.

Bliege Bird, Rebecca, "Cooperation and Conflict: the Behavioral Ecology of the Sexual Division of Labor," *Evolutionary Anthropology* 8 (1999):65–75.

Bliege Bird, Rebecca L., Eric A. Smith, and Douglas W. Bird, "The Hunting Handicap: Costly Signaling in Human Foraging Strategies," *Behavioral Ecology and Sociobiology* 50 (2001):9–19.

Boehm, Christopher, "Segmentary warfare and Management of Conflict: a Comparison of East African Chimpanzees and Patrilineal-patrilocal Humans," in A. Harcourt and Frans de Waal (eds.) *Coalitions and Alliances in Humans and Other Animals* (Oxford: Oxford University Press, 1992) pp. 137–173.

Boone, James L., "The Evolution of Magnanimity: When is it Better to Give than to Receive?," *Human Nature* 9 (1998):1–21.

Bowles, Samuel, "Individual Interactions, Group Conflicts, and the Evolution of Preferences," in Steven Durlauf and Peyton Young (eds.) *Social Dynamics* (Cambridge, MA: MIT Press, 2001) pp. ??–??

Boyd, Robert and Peter J. Richerson, *Culture and the Evolutionary Process* (Chicago: University of Chicago Press, 1985).

— and — , "The Evolution of Cooperation," *Journal of Theoretical Biology* 132 (1988):337–356.

— and — , "Group Selection among Alternative Evolutionarily Stable Strategies," *Journal of Theoretical Biology* 145 (1990):331–342.

— and — , "Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizeable Groups," *Ethology and Sociobiology* 113 (1992):171–195.

Brown, Jerram L., "Cooperation—A Biologist's Dilemma," *Advances in the Study of Behavior* 13 (1983):1–37.

Cavalli-Sforza, Luigi L. and Marcus W. Feldman, *Cultural Transmission and Evolution* (Princeton, NJ: Princeton University Press, 1981).

Clutton-Brock, T. H. and G. A. Parker, "Punishment in Animal Societies," *Nature* 373 (1995):58–60.

Connor, Richard C., "The Benefits of Mutualism: a Conceptual Framework," *Biological Reviews* 70 (1995):427–457.

Dieckmann, Ulf, "Can Adaptive Dynamics Invade," *Trends in Ecology and Evolution* 12 (1997):128–131.

Dugatkin, Lee Alan, “Tit for Tat, Byproduct Mutualism and Predator Inspection: A reply to Connor,” *Animal Behavior* 51 (1996):455–457.

— , *Cooperation among Animals* (New York: Oxford University Press, 1997).

— and Michael Mesterton-Gibbons, “Cooperation Among Unrelated Individuals: Reciprocal Altruism, Byproduct Mutualism, and Group Selection in Fishes,” *Biosystems* 37 (1996):19–30.

Fehr, Ernst and Simon Gächter, “Cooperation and Punishment,” *American Economic Review* 90,4 (September 2000).

Frank, Steven A, “Mutual Policing and Repression of Competition in the Evolution of Cooperative Groups,” *Nature* 377 (October 1995):520–522.

Getty, Thomas, “Handicap Signaling: When Fecundity and Viability do not Add Up,” *Animal Behavior* 56 (1998):127–130.

Gintis, Herbert, *Game Theory Evolving* (Princeton, NJ: Princeton University Press, 2000).

— , “Strong Reciprocity and Human Sociality,” *Journal of Theoretical Biology* 206 (2000):169–179.

Godfray, J. C. J., “Signaling of Need by Offspring to their Parents,” *Nature* 352 (1991):328–330.

Grafen, Alan, “Biological Signals as Handicaps,” *Journal of Theoretical Biology* 144 (1990):517–546.

— , “Sexual Selection Unhandicapped by the Fisher Process,” *Journal of Theoretical Biology* 144 (1990):473–516.

Gurven, Michael, Wesley Allen-Arave, Kim Hill, and A. Magdalena Hurtado, “’It’s a Wonderful Life’: Signaling Generosity Among the Ache of Paraguay,” *Evolution and Human Behavior* 21 (2000):263–282.

Hardin, Russell, *Collective Action* (Baltimore: Johns Hopkins University Press, 1982).

Harsanyi, John C., “Games with Incomplete Information Played by Bayesian Players, Parts I, II, and III,” *Behavioral Science* 14 (1967):159–182, 320–334, 486–502.

Hawkes, Kristen, “Why Hunter-Gatherers Work: An Ancient Version of the Problem of Public Goods,” *Current Anthropology* 34,4 (1993):341–361.

Hawkes, Kristin, J. F. O’Connell, and N. G. Blurton Jones, “Hadza Meat Sharing,” *Evolution and Human Behavior* 22 (2001):113–142.

Hofbauer, Josef and Karl Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge, UK: Cambridge University Press, 1998).

Iwasa, Yoh and Andrew Pomiankowski, “Good Parent and Good Genes Models of Handicap Evolution,” *Journal of Theoretical Biology* 200 (1999):97–109.

Johnstone, Rufus A., “Sexual Selection, Honest Advertisement and the Handicap Principle,” *Biological Reviews* 70 (1995):1–65.

— , “The Evolution of Animal Signals,” in John R. Krebs and Nicholas B. Davies (eds.) *Behavioral Ecology: An Evolutionary Approach (4th ed.)* (Oxford: Blackwell, 1997) pp. 155–178.

— , “Signaling of Need, Sibling Competition, and the Cost of Honest Signaling,” *Proceedings of the National Academy of Sciences* 96,22 (1999):12644–12649.

Kaplan, Hillard and Kim Hill, “Hunting Ability and Reproductive Success among Male Ache Foragers: Preliminary Results,” *Current Anthropology* 26,1 (1985):131–133.

Lachmann, Michael and Carl T. Bergstrom, “Signalling among Relatives, II. Beyond the Tower of Babel,” *Theoretical Population Biology* 54 (1998):146–160.

Leimar, O. and P. Hammerstein, “Evolution of Cooperation through Indirect Reciprocity,” *Proc. Royal. Soc. Lond. B (*2001):745–753.

Marlowe, Frank, “Showoffs or Providers: The Parenting Effort of Hadza Men,” *Evolution and Human Behavior* 20,6 (1999):391–404.

Maynard Smith, John, “Honest Signalling: The Philip Sidney Game,” *Animal Behavior* 42 (1991):1034–1035.

Mesterton-Gibbons, Michael and Lee Alan Dugatkin, “Cooperation and the Prisoner’s Dilemma: Toward Testable Models of Mutualism Versus Reciprocity,” *Animal Behavior* 54 (1997):1551–557.

Michod, Richard E., “Cooperation and Conflict in the Evolution of Individuality. 1. The Multilevel Selection of the Organism,” *American Naturalist* 149 (April 1997):607–645.

Milinski, Manfred, “Byproduct Mutualism, Tit-for-tat Reciprocity and Cooperative Predator Inspection,” *Animal Behavior* 51 (1996):458–461.

Mitani, John C. and David P. Watts, “Why Do Chimpanzees Hunt and Share Meat?,” *Animal Behaviour* 61 (2001):915–924.

Nowak, Martin A. and Karl Sigmund, “Evolution of Indirect Reciprocity by Image Scoring,” *Nature* 393 (1998):573–577.

Pusey, Anne E. and Craig Packer, “The Ecology of Relationships,” in John R. Krebs and Nicholas B. Davies (eds.) *Behavioral Ecology: An Evolutionary Approach,* 4th ed. (Oxford: Blackwell Science, 1997) pp. 254–283.

Roberts, Gilbert, “Competitive Altruism: From Reciprocity to the Handicap Principle,” *Proceedings of the Royal Society of London B* 265 (1998):427–431.

Ruyle, Eugene E., “Slavery, Surplus, and Stratification on the Nortwest Coast: the Ethnoenergetics of an Incipient Stratification System,” *American Anthropologist* 14 (1973):603–631.

Schelling, Thomas C., *Choice and Consequence* (Cambridge, MA: Harvard University Press, 1978).

Selten, Reinhard, "A Note on Evolutionarily Stable Strategies in Asymmetric Animal Conflicts," *Journal of Theoretical Biology* 84 (1980):93–101.

Siller, Steven, "The Epistatic Handicap Principle Does Work," *Journal of Theoretical Biology* 191 (1998):141–161.

Smith, Eric Alden and Rebecca L. Bliege Bird, "Turtle Hunting and Tombstone Opening: Public Generosity as Costly Signaling," *Evolution and Human Behavior* 21,4 (2000):245–261.

Sober, Elliot and David Sloan Wilson, *Unto Others: The Evolution and Psychology of Unselfish Behavior* (Cambridge, MA: Harvard University Press, 1998).

Sosis, Richard, "Costly Signaling and Torch Fishing on Ifaluk Atoll," *Evolution and Human Behavior* 21 (2000).

Stanford, C. B., J. Wallis, E. Mpongo, and J. Goodall, "Hunting Decisions in Wild Chimpanzees," *Behavior* 131 (1994):1–20.

Stanford, Craig B., *The Hunting Apes* (Princeton, NJ: Princeton University Press, 1999).

Taylor, P. and L. Jonker, "Evolutionarily Stable Strategies and Game Dynamics," *Mathematical Biosciences* 40 (1978):145–156.

Trivers, R. L., "The Evolution of Reciprocal Altruism," *Quarterly Review of Biology* 46 (1971):35–57.

Veblen, Thorstein, *The Theory of the Leisure Class* (New York: Macmillan, 1899).

Wedekind, Claus and Manfred Milinski, "Cooperation Through Image Scoring in Humans," *Science* 289 (5 May 2000):850–852.

Wilson, David Sloan, "Structure Demes and the Evolution of Group-advantageous Traits," *American Naturalist* 111 (1977):157–185.

Wrangham, R. W., "An Ecological Model of Female-bonded Primate Groups," *Behavior* 75 (1980):262–300.

Wright, Jonathan, "Altruism as Signal: Zahavi's Alternative to Kin Selection and Reciprocity," *Journal of Avian Biology* 30 (1999):108–115.

Zahavi, Amotz, "The Cost of Honesty (Further Remarks on the Handicap Principle)," *journal of Theoretical Biology* 67 (1977):603–605.

— , "Reliability in Communication Systems and the Evolution of Altruism," in B. Stonehouse and C. M. Perrins (eds.) *Evolutionary Ecology* (London: Macmillan Press, 1977) pp. 253–259.

— , "Altruism as Handicap: the Limitations of Kin Selection and Reciprocity," *Journal of Avian Biology* 26 (1995):1–3.

— and Avishay Zahavi, *The Handicap Principle: A Missing Piece of Darwin's Puzzle* (New York: Oxford University Press, 1997).

e\Papers\Eric Smith Paper\Costly Signaling and Reciprocity July 17, 2001