Adaptive Behavior

http://adb.sagepub.com

Cooperation, Conflict and the Evolution of Communication

Jason Noble

Adaptive Behavior 1999; 7; 349

DOI: 10.1177/105971239900700308

The online version of this article can be found at:
http://adb.sagepub.com/cgi/content/abstract/7/3-4/349

Published by:
SAGE Publications
http://www.sagepublications.com

On behalf of:
ISAB
International Society of Adaptive Behavior

Additional services and information for *Adaptive Behavior* can be found at:

Email Alerts: http://adb.sagepub.com/cgi/alerts

Subscriptions: http://adb.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations (this article cites 16 articles hosted on the
SAGE Journals Online and HighWire Press platforms):
http://adb.sagepub.com/cgi/content/refs/7/3-4/349

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

Adaptive Behavior 1999, 7(3/4), 349-370

Cooperation, Conflict and the Evolution of Communication

Jason Noble
Center for Adaptive Behavior and Cognition
Max Planck Institute for Human Development

This paper presents a general model that covers signaling with and without conflicts of interest between signalers and receivers. Krebs and Dawkins (1984) argued that a conflict of interests will lead to an evolutionary arms race between manipulative signalers and sceptical receivers, resulting in ever more costly signals; whereas common interests will lead to cheap signals or "conspiratorial whispers." Previous simulation models of the evolution of communication have usually assumed either cooperative or competitive contexts. Simple game-theoretic and evolutionary simulation models are presented; they suggest that signaling will evolve only if it is in the interests of both parties. In a model where signalers may inform receivers as to the value of a binary random variable, if signaling is favored at all, then signalers will always use the cheapest and the second cheapest signal available. Costly signaling arms races do not get started. A more complex evolutionary simulation is described, featuring continuously variable signal strengths and reception thresholds. As the congruence of interests between the parties becomes more clear-cut, successively cheaper signals are observed. The findings support a modified version of Krebs and Dawkins's argument. Several variations on the continuous-signaling model are explored.

Keywords: behavioral ecology, communication, competition, coevolution, cooperation, signaling.

1 THE PROBLEM OF COMMUNICATION

Communication is an important aspect of the social behavior of animals. Mammals and birds give alarm calls to warn conspecifics of approaching predators; ants and termites use signals to recruit others to the defence of the nest; bees and primates inform others in their colony or group of the discovery of new food sources. Such communication systems typically have obvious benefits. Consider the alarm calls of vervet monkeys, described by Seyfarth, Cheney, and Marler (1980). Vervets call when they detect one of three (or sometimes four) distinct categories of predator, and the response of the monkeys who hear the call is appropriate: the leopard alarm, for instance, sends hearers scrambling into the relative safety of nearby trees. Surely the adaptive function of this communication system is to help the monkeys avoid predation? At one time, when group selection as a force in evolution was held in higher regard than it is today, this explanation might have been uncritically accepted. However, the orthodox position in evolutionary biology (Williams, 1966; Dawkins, 1976; Maynard Smith, 1993) now tells us that animals are best understood as products of their selfish genes: animals do not do things for the good of the group or the species, but in order to propagate copies of their genetic material. From this viewpoint, avoiding predation is likely only to be the function of the response behavior. The function of the signaling behavior is not so obvious: why should a monkey that has spotted an approaching leopard warn its conspecifics? Giving the alarm signal may well increase the risk to the signaler, by drawing the leopard's attention to itself.

This problem applies to many social signals such as alarm and food calls. It is easy to see where the benefit lies for receivers of the signal; being informed of the approach of a predator or the location of food is clearly useful. It is not so easy, however, to determine why the signaler should share the relevant information. In many contexts there will either be no benefit in doing so, or, more likely, costs involved. These costs may be due to, for example, energy expenditure in the production of the signal, an increase in personal risk for

Copyright 1999 International Society for
Adaptive Behavior

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

the signaler, or the loss of food that might have been consumed alone. There is thus a degree of altruism in such signaling.

The problem of accounting for honesty becomes even more acute when we consider communication that occurs—or appears to occur—despite an explicit conflict of interests between signalers and receivers. For example, in aggressive or territorial signals, each animal would prefer that the other respond by retreating, and might be expected to exaggerate signals of strength or willingness to attack. In many sexual advertisement signals, it is in the interests of the average male to convince any female he meets to copulate with him, but it is in the average female's interests to be difficult to persuade, and to mate only with the highest-quality males. In these cases, the potential benefits for a dishonest strategist can be great.

Even in the apparently cooperative context of alarm and food calls, what prevents the invasion of free-riders who gain the benefit of others' honest signals, but do not pay the costs of honesty themselves? How can honest signaling be an evolutionarily stable strategy (ESS)? Furthermore, how might such communication have evolved in the first place—why, against an initial background of non-communication, would the first proto-signalers have been selected for their behavior?

Reciprocal altruism (Trivers, 1971), an arrangement in which one animal bears a cost in order that another may gain a (greater) benefit and later has the favor returned, provides a possible answer. Simulation work such as Axelrod and Hamilton (1981) and Axelrod (1984) has shown that reciprocal altruism can generate cooperative behavior in a situation where the immediate short-term interests of the participants conflict. However, reciprocal altruism requires relatively sophisticated cognitive machinery, in terms of the ability to identify others and to remember their record of cooperation or defection. It is reasonable to postulate such cognitive machinery in, for example, primates, but it is not likely to be common in the rest of the animal world. Reciprocal altruism may well be implicated in such primate aberrations as human language, but the current paper seeks more general mechanisms.

Kin selection (Hamilton, 1964) may also be a factor in explaining the existence of stable natural communication systems. Kin selection refers to the idea that an animal has two ways of ensuring the survival of its genes: firstly, by direct reproduction, and secondly by assisting relatives in their reproductive efforts, thereby promoting the survival of copies of at least some of its genes. In the case of alarm calls, even though a calling animal might suffer a personal cost due to increased predation risk, so long as the call tends to benefit its genetic relatives to a sufficient degree, calling behavior will be evolutionarily stable. In Sherman's (1977) investigation of alarm calls in Belding's ground squirrel, calls were most often given by females who lived in groups of related individuals; kin selection was found to be the most likely hypothesis to explain the behavior.

Kin selection can thus explain apparently altruistic communication among groups of close relatives. However, communication systems in the animal world are not limited to those species that live in groups of closely related individuals. What other mechanisms, apart from reciprocal altruism and kin selection, might foster the evolution of communication despite potential conflicts of interest between signalers and receivers? Under what circumstances should we expect would-be communicators to overcome the perennial problems of cheating, lying and bluffing? A third possible evolutionary mechanism that could lead to honesty in a signaling system is the handicap principle (Zahavi, 1975, 1987). This is the idea that honesty can be maintained if the signals are costly in a particular way. However, the handicap principle has recently received overwhelming theoretical attention (see for example Grafen, 1990; Iwasa, Pomiankowski, & Nee, 1991; Hurd, 1995; Bullock, 1997) and will only be treated tangentially here.

1.1 Manipulative and cooperative signaling

Krebs and Dawkins (1984) provide another possible answer, and in so doing challenge the default notion that animal communication is about information transmission. They suggest that propaganda and advertising make better metaphors for animal communication than does the cooperative use of language to share information. They view signaling as a competitive affair involving mind-reading and manipulation. Mind-reading consists of one animal exploiting tell-tale predictors about the future behavior of another, e.g., a dog noticing the bared teeth of an opponent,

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

concluding that it is about to attack, and fleeing in order to avoid injury. Manipulation is what happens when those being mind-read fight back, influencing the behavior of the mind-readers to their own advantage. For example, a dog could bare its teeth despite not having the strength or inclination to attack, and thus scare off its mind-reading opponent. The authors predict evolutionary arms races between manipulative, exploitative signalers and sceptical receivers, as "selection will act simultaneously to increase the power of manipulators and to increase resistance to it" (p. 390). The result will be increasingly costly signals. In our example, the original subtle teeth-baring signal might become more and more elaborate, and be accompanied by growling, hair-bristling, staring, etc.

Krebs and Dawkins admit, however, that not all interactions are competitive in nature. There are some situations in which it is to the receiver's advantage to be manipulated by the signaler. For example, a pack-hunting predator may attempt to recruit a conspecific in order to bring down prey too large for either to tackle alone. Foraging bees, on returning to the hive, may indicate to their closely related hive-mates the direction and distance to a source of nectar. In these cases the receiver's compliance is to the benefit of both parties, i.e., there exists the possibility of cooperation. Krebs and Dawkins argue that when the two parties share a common interest in this way, then a different kind of signal coevolution will result. Specifically, there will be selection for signals that are as energetically cheap as possible while still being detectable; Krebs and Dawkins suggest the phrase "conspiratorial whispers" to describe these signals. Rather than signalers needing to be more and more extravagant in their attempts to persuade receivers, the opposite process occurs: receivers are eager to be persuaded, and selection will favor subtle signaling and low response thresholds. An implication is that the louder and costlier signaling displays of the animal world—such as roaring contests in red deer or male plumage in birds of paradise—may have been overrepresented in studies of animal communication simply because they are obvious to human observers. There may be a great deal of conspiratorial, cooperative signaling going on that is too subtle for us to notice.

Krebs and Dawkins's argument has been influential but no formal justification or model of it exists. One goal of the current paper is to test their prediction that evolved signals will necessarily be more costly when there is a conflict of interests than when the participants have common interests. In order to do so, it will first be necessary to determine whether communication should be expected at all when signalers and receivers have a genuine conflict of interests.

1.2 Simulation models of communication

The current paper uses both game-theoretic and evolutionary simulation models as tools for investigating communication in cooperative and competitive contexts. Previous models of the evolution of communication in the simulation of adaptive behavior (SAB) and artificial life (AL) literature have often been constructed such that honest signaling is in the interests of both signalers and receivers—any communication systems that evolve can therefore be described as cooperative. For example, Werner and Dyer (1992) postulated blind, mobile males and sighted, immobile females: The evolution of a signaling system was in the interests of both parties as it allowed mating to take place at better-than-chance frequencies. In MacLennan and Burghardt's (1994) model, signalers and receivers were rewarded if and only if they engaged in successful communicative interactions.

Other SAB/AL models (Ackley & Littman, 1994; Oliphant, 1996) have looked at the special case where communication would benefit receivers, but the potential signalers are indifferent. Oliphant argues that this is a good way to model the evolution of alarm calls: It captures the idea that the potential signaler already knows about the danger of the approaching predator, and tests the stability of a strategy of sharing that information. In fact, the models suggest that signaling will not evolve in these cases unless a mechanism such as reciprocal altruism or kin selection[1] is in place. Note that such mechanisms have no mystical effect: they simply shift the expected long-term inclusive-fitness payoffs for particular strategies such that communication is mutually beneficial.

Finally, some SAB/AL work considers the evolution of communication in situations where the two parties appear to have conflicting interests. Wheeler and de Bourcier (1995) modeled aggressive territorial signaling. Bullock (1997) constructed a general

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

model in which signalers of varying degrees of quality solicited receivers for a favorable response; receivers were rewarded for responding positively only to high-quality signalers. A conclusion drawn in both studies was that if signals were sufficiently costly (e.g., long, elaborate tails or energetic ritual displays) then reliable communication could evolve and persist over time—this is in line with Zahavi's handicap principle. Bullock made the more specific prediction that in order for communication to be stable, the net cost of signaling must be lower for higher-quality signalers (see also Grafen, 1990). However, it could be argued that such differential signal costs effectively render honest signaling mutually beneficial. We will return to this notion below.

A second goal of the current paper is to position previous SAB/AL work in an overarching theoretical context. To this end some general models of the evolution of simple signaling systems will be presented; the models will cover situations with and without a conflict of interests between the two interacting agents.

1.3 Conflicts of interest

The first requirement in constructing a general model of communication is a classification scheme for determining when a conflict of interests exists between signalers and receivers—Figure 1 shows such a scheme, adapted from Hamilton (1964). Assume that a successful instance of communication in a particular scenario has fitness implications for both participants. The fitness effect on signalers, P_S, and the fitness effect on receivers, P_R, together define a point on the plane in Figure 1. For example, consider a hypothetical food call, by which one animal alerts another to the presence of a rich but limited food source. By calling and thus sharing the food, the signaler incurs a fitness cost; by responding to the call, the receiver benefits through obtaining food it would otherwise have missed. Thus, the call would be located in the "altruism" quadrant. The situations modeled by Ackley and Littman (1994) and Oliphant (1996), where receivers benefit but signalers are ambivalent, can be thought of as points on the positive vertical axis, i.e., where $P_S = 0$ and $P_R > 0$.

Conflicts of interest can be defined as interactions in which natural selection favors different outcomes for each participant (Trivers, 1974), or in which participants place the possible outcomes in a different rank order (Maynard Smith & Harper, 1995). Conflicts of interest therefore exist when P_S and P_R are of opposite sign, i.e., in the upper-left and lower-right quadrants. Selection will, by definition, favor actions that have positive fitness effects. In the upper-left and lower-right quadrants, one agent but not the other will be selected to participate in the communication system: their interests conflict. The "spite" quadrant does *not* represent a conflict of interests because agents will be mutually selected not to communicate.

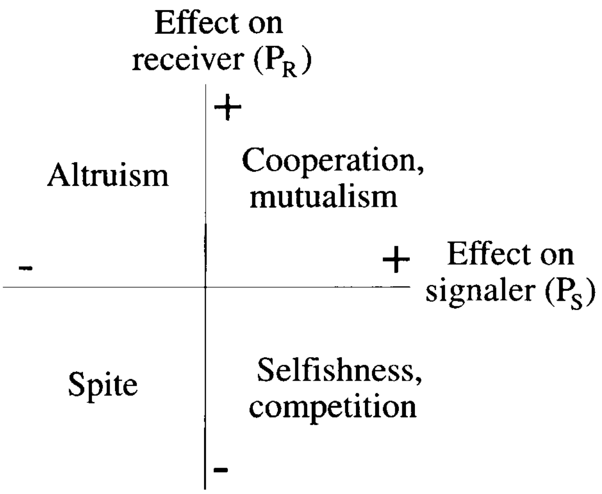


Figure 1. Possible communication scenarios classified by their effects on the fitness of each participant.

If the specified fitness effects of participating in a communicative interaction are truly *net* values, and already include such factors as the cost of signaling and the cost of making a response (as well as inclusive fitness considerations and costs due to exploitation of the signal by predators, etc.), then predicting the evolution of the communication system is trivial. Reliable communication requires, on average, honest signalers and trusting receivers, and thus will only develop when $P_S > 0$ and $P_R > 0$, i.e., when both agents are selected to participate. However, real animals sometimes communicate despite apparent conflicts of interest (Hinde, 1981). Recent models (Grafen, 1990; Bullock, 1997) have established that, in certain situations where communication would otherwise be unstable, increasing the production costs of the signal can lead to a prediction of evolutionarily stable signaling. Therefore, in the current model, P_S

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

and P_R refer to gross fitness effects before the specific costs of producing the signal, C_S, and making the response, C_R, have been taken into account.

2 A SIMPLE SIGNALING GAME

If the signaling interaction is to involve information transmission, and allow for the possibilities of deception and manipulation, it must be modeled as a game of imperfect information, in which the signaler knows something that the receiver does not. Some theorists might find this contentious: after all, Krebs and Dawkins (1984) have characterized animal communication as being more about persuasion than information, and Di Paolo (1997) has argued that the term “communication” should not be limited to situations of information exchange but defined more broadly as coordinated action. However, alarm and food calls clearly involve one animal with privileged access to an item of information about the world “deciding” whether or not to share its knowledge with others. Di Paolo may well be right about the boundaries of communication being wider than they are commonly thought to be, but that does not mean that an orthodox situation of possible information transmission fails to qualify.

Figure 2 shows the extended form of a simple action-response game that captures the structure of the alarm- or food-call context, and arguably other contexts besides. The game begins with a chance move (the central square) in which some state is randomly determined to be either “high” or “low.” The signaler has access to this state, and we can suppose that it represents either a feature of the environment that only the signaler has detected (e.g., noticing an approaching predator), or a hidden internal state of the signaler (e.g., ovulation). Based on this state, the signaler (player I) must decide whether or not to send an arbitrary signal of cost C_S. The receiver (player II) is ignorant of the hidden state and only knows whether or not a signal was sent—the dashed rectangles show the receiver’s information sets. The receiver can respond either positively, i.e., perform some action appropriate to the high state, or negatively, i.e., not respond at all. Positive responses incur a cost, C_R. If and only if the hidden state is high, a positive

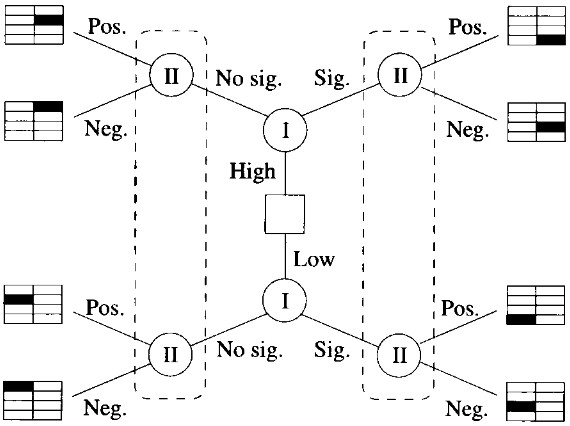


Figure 2. Extended form of the simple signaling game. Chart icons index payoffs in Table 1.

	State of environment	
	Low	High
No signal		
Neg. response	0,0	0,0
Pos. response	0, $-C_R$	P_S, P_R-C_R
Signal		
Neg. response	$-C_S$, 0	$-C_S$, 0
Pos. response	$-C_S$, $-C_R$	P_S-C_S, P_R-C_R

Table 1. Payoff matrix for the simple game. Entries in the table represent the payoff to the sender and receiver respectively.

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

response results in the payoffs P_S and P_R to the signaler and receiver respectively. Table 1 specifies the payoff matrix.

Hurd (1995), Oliphant (1996), and Bullock (1997) used similar games with different payoff structures. In each of these earlier games, the receiver was explicitly rewarded for accuracy in determining the hidden state. In contrast, in the current game accuracy is not a goal of the receiver per se; the receiver simply wants to maximize its average payoff. Depending on the precise payoff values, the best way to achieve that might be to respond in a blanket way, i.e., responding negatively or positively whatever the signal. This is meant to reflect the fact that receivers in natural contexts can presumably opt out of the communication system if it is to their advantage to do so; there is no force compelling them to pay attention to the signaler.

The game models a range of possible communicative interactions. For example, suppose that the high state represents the signaler's discovery of food. Sending a signal might involve making a characteristic sound, while not sending a signal is to remain silent. For the receiver, a positive response means approaching the signaler and sharing the food, whereas a negative response means doing nothing. Various possibilities exist besides honest signaling of the high state: The receiver might *always* approach the signaler in the hope of obtaining food, regardless of whether a signal was sent. The signaler might be uninformative and never signal, or only signal when food was not present. One important feature of the game is that the signaler is ambivalent about the receiver's response in the low state—in terms of the example, this represents the assumption that when no food has been discovered, the signaling animal does not care about whether the receiver approaches or not.

The strategies favored at any one time will depend on the relative values of P_S, P_R, C_S and C_R, as well as on what the other members of the population are doing. (Another parameter of interest is the relative frequency of high and low states; in the models presented here each state occurred 50% of the time.) Allowing the base fitness effects P_S and P_R to vary across positive and negative values will allow the payoff space of Figure 1 to be explored, and thus determine whether changes in signal and response cost can produce stable signaling in situations that would otherwise involve conflicts of interest. This will be a first step towards assessing Krebs and Dawkins's conspiratorial whispers theory.

2.1 Stable strategies in the simple game

A signaling strategy in the simple game specifies whether to respond with no signal (NS) or a signal (Sig) to low and high states respectively. Likewise, a response strategy specifies whether to respond negatively (Neg) or positively (Pos) when faced with no signal and when faced with a signal. A complete strategy is the conjunction of a signaling and a response strategy; e.g., (NS/NS, Pos/Pos) is the strategy that specifies never signaling and always responding positively.

The strategy (NS/Sig, Neg/Pos) specifies signaling only in the high state, and responding positively only to signals—call this the "honest and trusting" strategy. Evolutionary stability depends on a strategy being the best response to itself; i.e., a strategy must be uninvadable in order to be an ESS. Honest and trusting players meeting each other can expect an average payoff per interaction of:

$$(P_S - C_S + P_R - C_R) / 4.$$

This will be higher than the expected payoff for any possible invading strategy (i.e., honesty and trust will be an ESS) if:

$$P_S > C_S > 0$$
$$P_R > C_R > 0.$$

That is, honest signaling is stable if the costs of signaling and responding are both positive, and if the payoffs in each case outweigh the costs. The requirement that P_S and P_R must both be positive means that the honest strategy is only expected to be stable when the interests of the parties do not conflict: positive values of P_S and P_R place the interaction in the upper right "mutualism" quadrant of Figure 1.

Of the 16 possible strategy pairs, there are three besides the honest strategy that involve the transmission of information, in that the receiver responds differently to different hidden states. None of these three strategy pairs are ESSs if C_S and C_R are both positive; these two values represent energetic costs and so cannot sensibly be negative. If $C_S = 0$, i.e., if giving a signal is of negligible cost, then the reverse honesty strategy (Sig/NS, Pos/Neg) can be stable, although P_S and

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

P_R must still be positive. It is also worth noting that a population consisting entirely of individuals playing (NS/NS, Pos/Pos) or (NS/NS, Pos/Neg), both non-signaling strategies where the receiver always responds positively, cannot be invaded by any other strategy if the payoff to the receiver is large enough, i.e., if:

$$P_S > C_S > 0$$
$$P_R > 2C_R > 0.$$

The analysis indicates that while the cost of signaling plays some role in stabilizing the honest strategy, there are no circumstances in which stable communication is predicted when a conflict of interests exists. This is despite the fact that we have separated the costs of signaling and responding from the base fitness payoffs of a communicative interaction.

2.2 Evolutionary simulation model

Game theory is limited to describing equilibria; an evolutionary simulation model of the simple game was also constructed in order to determine whether communicative behavior might sometimes be found outside the range of identified ESSs.

A straightforward genetic algorithm (GA) was used. Each individual could play both signaling and receiving roles; a strategy pair was specified by a four-bit genotype as shown in Table 2. The population size was 100, the mutation rate was 0.01 per locus, and, due to the trivially small genome, crossover was not used. Each generation, 500 games were played between randomly selected opponents. An agent could therefore expect to play five games as a signaler and five as a receiver. The agent's fitness score was the total payoff from these games. For breeding purposes, the fitness scores were normalized by subtracting the minimum score from each. Proportionate selection was then applied to the normalized scores. The genetic algorithm was run in this manner for 500 generations. In the results presented below, the games played in the final, i.e., 500th, generation have been used as a snapshot of the evolved signaling strategies.

An attempt was made to investigate evolutionary dynamics, in that the initial populations were not determined randomly but started as either "honest" or "non-signaling." Honest initial populations were made up entirely of individuals who played the honest and trusting strategy, i.e., a genome of "0101." Non-signaling populations underwent 100 generations of preliminary evolution in which their receiving strategies were free to evolve but their signaling strategies were clamped at "00," i.e., no signaling. For each class of initial conditions, a simulation run was performed for all combinations of integer values of P_S and P_R between -5 and +5, making 121 runs in all. Each run was repeated 25 times with different random seeds. The values of C_S and C_R were fixed at 1.

Communication was indexed by cross-tabulating the hidden state value with the receiver's response and calculating a chi-squared statistic. The receiver has no direct access to the hidden state, so any reliable correspondence between state and response indicates that information has been transmitted and acted upon. Values of the χ^2 statistic close to zero indicate no communication, and values close to the maximum (in this case $\chi^2_{max} = 500$, due to the 500 games played in the final, snapshot generation) indicate near-perfect communication.

Figure 3 shows the average values of the communication index for honest initial conditions. Seeding the population with honesty tests the stability of honest

	Bit value	
	0	**1**
If low state	No signal	Signal
If high state	No signal	Signal
Response to no signal	Negative	Positive
Response to signal	Negative	Positive

Table 2. Genetic specification of strategies.

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

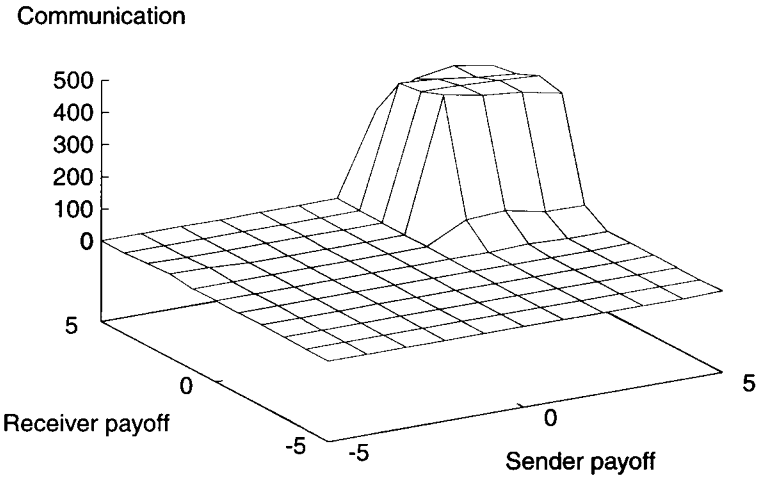


Figure 3. Mean communication index by P_S and P_R; honest initial conditions. Each point is a mean calculated over 25 runs. Mean standard error = 2.96.

signaling given a particular payoff pair, much as a game-theoretic analysis does. The results are compatible with the conditions outlined in the previous section: Honesty is stable when the payoffs to signaling and receiving are positive and greater than the respective costs. However, there is some suggestion of intermittent or imperfect communication when $P_R = C_R = 1$, indicating that ambivalent receivers may occasionally cooperate.

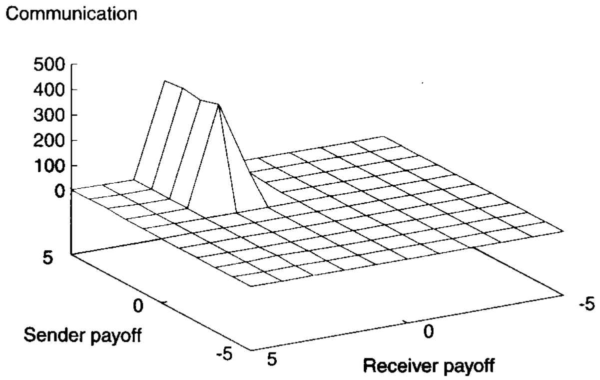


Figure 4. Mean communication index by P_S and P_R; non-signaling initial conditions. Each point is a mean calculated over 25 runs. Mean standard error = 1.67. Graph rotated for clarity—cooperative quadrant appears at top left.

Figure 4 shows the average values of the communication index for non-signaling initial conditions. Starting the GA with a non-signaling population tests the likelihood that communication will emerge, given a particular payoff pair. Clearly the conditions for emergence and stability-once-present are not the same. If $P_S > 1$ and $P_R = 2$ communication develops but when $P_S > 1$ and $P_R > 2$ it does not.

In the latter region $P_R > 2\ C_R$ and the population remains at the non-signaling equilibrium described in Section 2.1. Despite the fact that communication would result in a higher average fitness, the high value of P_R keeps the receivers responding positively all the time, removing any incentive for the signalers to bother signaling. This response strategy could be called "blind optimism," as receivers always respond positively. It should be noted, however, that the condition $P_R > 2\ C_R$ is dependent on the 50% frequency of high states; if high states occurred 10% of the time for instance, then $P_R > 10\ C_R$ would be required to make blind optimism a stable strategy.

The difference in results between the two classes of initial conditions is interesting but should not obscure the fact that no communication was observed under conditions of conflicting interests. We must conclude that, at least in the simple model discussed so far, stable communication is only to be expected when it is

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

in the interests of both parties.

3 A GAME WITH CONTINUOUS SIGNAL COSTS

In the simple signaling game, signalers can choose between a costly signal or no signal at all. The model does not allow for a range of possible signals with differing costs, and in this respect it is unrealistic. It may be that Krebs and Dawkins's implicit prediction, that signaling can occur when a conflict of interests exists, is in fact true, but can only be demonstrated in a more complex game with a range of signal costs. The simple signaling game (see Figure 2) was therefore extended to incorporate signals of differing costs.

3.1 Stable strategies in the continuous-signal-cost game

In the extended game, the signaling player has three options: not signaling, which costs nothing; using the "soft" signal, which costs C_S, and using the "loud" signal, which costs $2C_S$. Strategies in the extended game require specifying the signal to give when the hidden state is low, the signal to give when it is high, and the response to give to each of no-signal, soft and loud. The two strategies representing conspiratorial whispers or cheap signaling are (NS/Soft, Neg/Pos/Pos) and (NS/Soft, Neg/Pos/Neg). Both strategies call for the soft signal to be used in the high state, and for positive responses to the soft signal; the strategies differ only in the response to loud signals. Neither of these strategies can strictly be considered an ESS on its own (because neutral drift can take the population from one to the other) but it can be shown that the set of all mixed strategies involving these two is an ESS under the familiar conditions:

$$P_S > C_S > 0$$
$$P_R > C_R > 0.$$

Costly signaling would involve the use of the loud signal for the high state, and either the soft signal or no signal to denote the low state, with a corresponding response strategy. None of the four strategies in this category can be an ESS. For example, (NS/Loud, Neg/Pos/Pos) cannot be an ESS assuming positive costs of signaling and responding. The similar strategy (NS/Loud, Neg/Neg/Pos) is almost stable if $P_S > 2C_S$, but can drift back to the previous strategy which can in turn be invaded by the cheap strategy (NS/Soft, Neg/Pos/Pos).

Analysis of the extended game indicates that if signaling is favored at all, then at equilibrium the signalers will always use the cheapest and the second-cheapest signal available (i.e., no signal and the soft signal). Extending the game by adding ever more costly signaling options, until we have approximated a continuous range of signal costs, does not alter this conclusion. None of the costly signaling strategies can even be an ESS, let alone support communication in the face of a conflict of interests. The possibility of expensive signaling arms races starts to look remote. However, it may be that an evolutionary simulation model will reveal signaling strategies that, while unstable in the long term, nevertheless lead to transient communication under conditions of conflicting interest.

3.2 Evolutionary simulation model

A second evolutionary simulation was constructed, in which the cost of signaling was continuously variable. Signaling strategies were represented by two positive real numbers C_{low} and C_{high}: the cost of the signals given in the low state and in the high state respectively. Response strategies were represented by a real-valued threshold T; positive responses were given to signals with costs greater than the receiver's threshold value. Note that threshold value could be negative, indicating a positive response to any signal.

A real-valued GA was used to simulate the evolution of strategies over time. Generally, the same parameters were used as in the previous simulation model, e.g., a population of 100. Mutation was necessarily a different matter: each real-valued gene in each newborn individual was always perturbed by a random gaussian value, $\mu = 0$, $\sigma = 0.05$. If a perturbation resulted in a negative cost value the result was replaced by zero. In addition, 1% of the time (i.e., a mutation rate of 0.01) a gene would be randomly set to a value between 0 and 5 for signal costs, or between -5 and +5 for the threshold value. This two-part mutation regime ensured that offspring were always slightly different from their parent, and occasionally very different.

The C_S parameter was no longer relevant, but C_R, the cost of responding, remained fixed at 1. Honest initial conditions were implemented by setting

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

C_{low} = 0, C_{high} = 1.0 and T = 0.5. Non-signaling initial conditions were implemented by setting T to a random gaussian (μ =0, σ =1) and then clamping C_{low} = C_{high} = 0 for 100 generations of preliminary evolution.

Figures 5 and 6 show the average values of the communication index for honest and non-signaling initial conditions respectively. The results are qualitatively similar to those of the discrete simulation model: communication occurs in both cases, but in a more limited range of the payoff space for non-signaling conditions. In neither case does communication occur outside the "cooperative" quadrant.

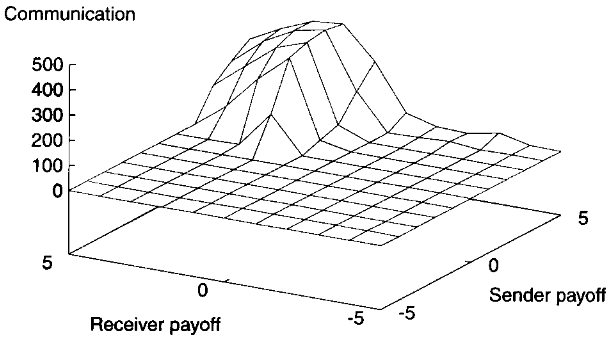


Figure 5. Mean communication index by P_S and P_R in the continuous simulation; honest initial conditions. Each point is a mean calculated over 25 runs. Mean standard error = 4.22. Graph rotated for clarity—cooperative quadrant appears at top.

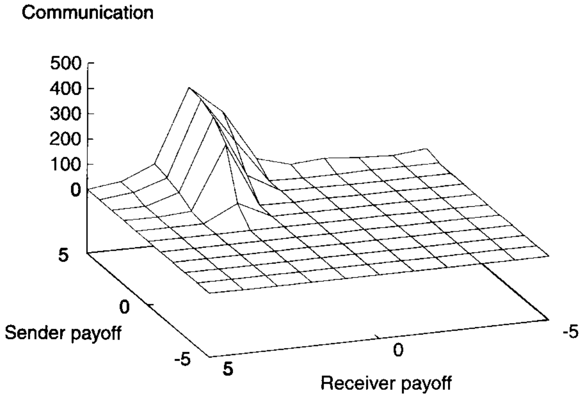


Figure 6. Mean communication index by P_S and P_R in the continuous simulation; non-signaling initial conditions. Each point is a mean calculated over 25 runs. Mean standard error = 3.61. Graph rotated for clarity—cooperative quadrant appears at top left.

However, there is some evidence that transient communication can occur when the conflict of interests between the two agents is not too extreme. For example, consider the payoff pair P_S = 5 and P_R = 0. This defines a point on the boundary between mutualism and selfishness, although when the constant cost of responding (C_R = 1) is taken into account, the net payoffs indicate that communication under these circumstances would be selfish (from the point of view of the signaler). Nevertheless, as Figure 7 shows, unstable communication evolves, even from non-signaling initial conditions.

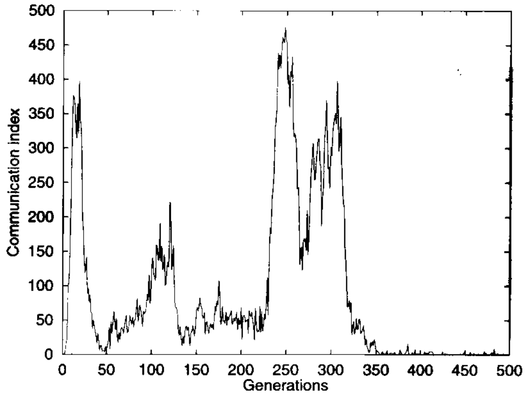


Figure 7. Mean communication index over plotted over generational time. A typical run with P_S = 5, P_R = 0, and non-signaling initial conditions.

The continuous model also allows investigation of the cost and threshold values over the payoff space. C_{low}, the cost of the signal given in response to the low state, always remained close to zero—this was unsurprising as signalers are ambivalent about the receiver's response to the low state. However, the value of C_{high} varied both inside and outside the region where communication was established: Figure 8 shows the mean values of C_{high} for honest initial conditions. The signals given in response to the high state are most costly when P_S, the payoff to the sender, is high and when the receiver's net payoff is marginal, i.e., $P_R \approx 1$. In order to study this effect more closely, additional simulation runs were performed, with P_S fixed at 5 and P_R varied between -5 and +5 in increments of 0.1. These runs can be thought of as exploring the cross section through P_S = 5 in Figure 8. Figure 9 shows the cross-sectional mean values of C_{high}. Note that the

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

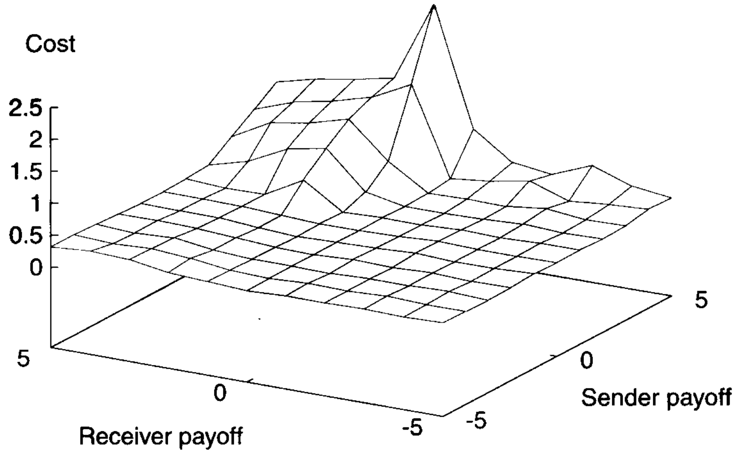


Figure 8. Mean cost of high-state signals by P_S and P_R; honest initial conditions. Each point is a mean calculated over 25 runs. Mean standard error = 0.032. Graph rotated for clarity—cooperative quadrant appears at top.

"energy" devoted to signaling is at a maximum around $P_R = 1$ and drops off as P_R increases—it can be seen from Figure 5 that $P_R = 1$ is approximately the point where significant communication is established. The same pattern was observed for non-signaling initial conditions (not shown for reasons of space).

The threshold values show corresponding variation. Figure 10 shows the mean value of T across the payoff space. The threshold values are typically very high (a "never respond" strategy) or very low (an "always respond" strategy), but in the region where communication evolved, receivers become progressively less demanding, i.e., T gets lower as P_R increases. Figure 11 shows the cross-sectional results for $P_S = 5$.

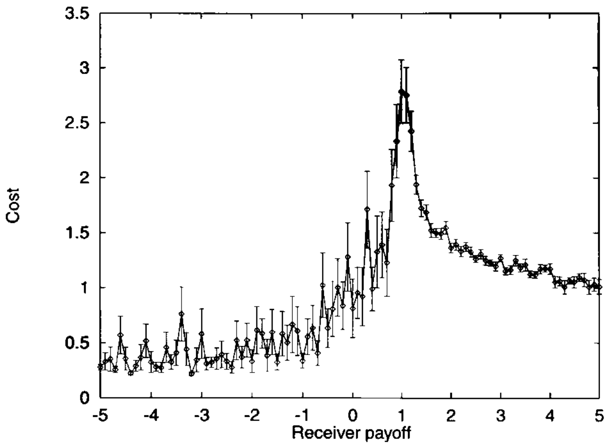


Figure 9. Cross-sectional means (±1 standard error) for high-state signal costs with $P_S = 5$; honest initial conditions. Each point is a mean calculated over 25 runs.

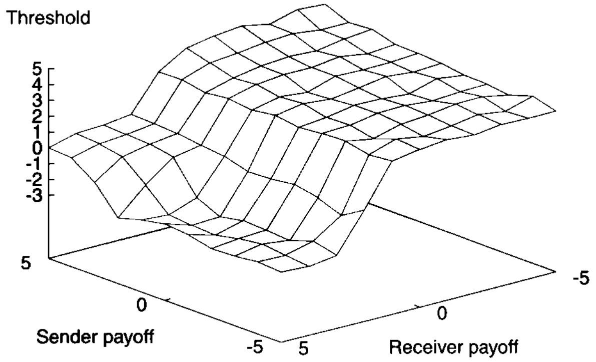


Figure 10. Mean threshold value by P_S and P_R; honest initial conditions. Each point is a mean calculated over 25 runs. Mean standard error = 0.19. Graph rotated for clarity—cooperative quadrant appears at top left.

Figure 12 plots the mean cost of high and low signals and the mean reception threshold all on one graph. This makes the relationship between costs and threshold clear: At approximately $P_R = 1$, the threshold falls to a level where the mean high-state signal will generate a positive response. As P_R increases, i.e., as the two players' payoffs approach each other, the signalers

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

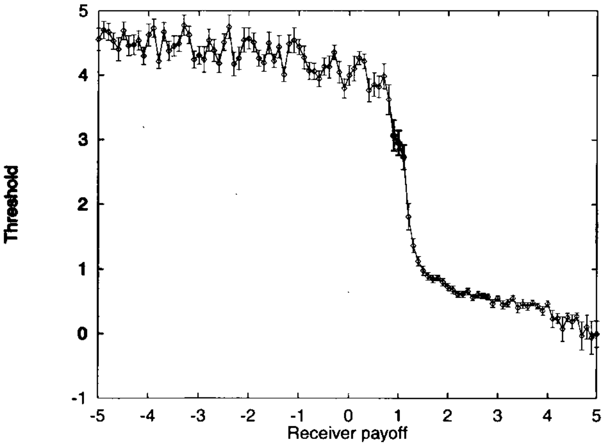


Figure 11. Cross-sectional mean threshold values (±1 standard error) with $P_S = 5$; honest initial conditions. Each point is a mean calculated over 25 runs.

become less extravagant and the receivers less "sceptical." This is *contra* the game-theoretic result of the previous section, which implies that when signals of varying costs are available, either the cheapest pair of signals will be used, or no signaling will occur—something like Figure 13 would be expected if the soft-loud signaling game accurately modeled the continuous case.

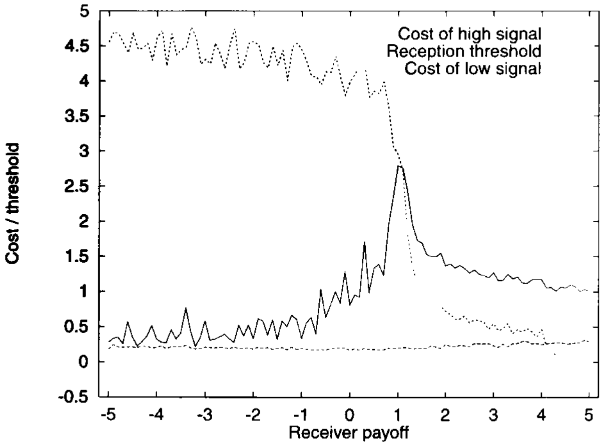


Figure 12. Cross-sectional means: cost of high and low signals, and reception threshold. $P_S = 5$, honest initial conditions. Each point is a mean calculated over 25 runs.

Note that the initial values of C_{high} and T under honest initial conditions were 1.0 and 0.5 respectively. For all but the highest values of P_R, C_{high} has increased on average over the 500-generation run. This rules out any explanation of the results of Figure 12 in terms of there having been insufficient evolutionary time for a cheaper signaling equilibrium to have been reached when the profit for receivers (P_R - C_R) was marginal. Evolution has taken the populations away from the cheap signaling solution.

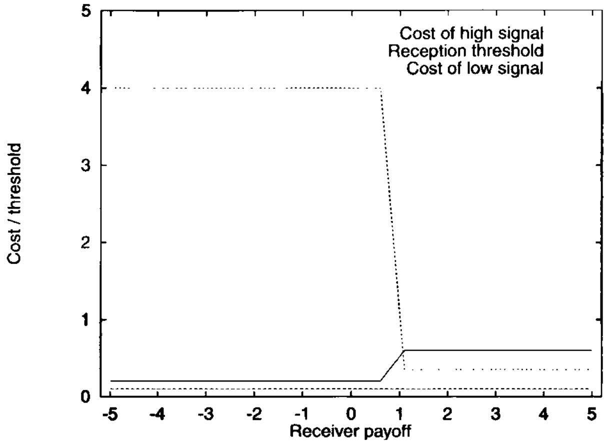


Figure 13. Approximate predicted results for Figure 12 according to discrete-cost game-theoretic model.

3.3 Discussion

In all of the models presented so far, stable communication evolved or was predicted to evolve only within the cooperative region of the signaler-receiver payoff space. This means that no signaling at all (costly or otherwise) was observed when the signaler and the receiver were experiencing a conflict of interests, apart from transitory communication on the boundaries of the cooperative region as shown in Figure 7.

The second game-theoretic model, in which discrete signals of varying costs are available, suggests that communication, if selected for, will involve the cheapest pair of signals available. However, the second simulation model, incorporating the more realistic assumption that signals can vary continuously in cost, implies that cheap signals will only be used when both parties stand to gain a high payoff from effective communication. When the net payoff to the receiver is marginal, evolved signals will be more costly than strictly necessary to convey the information. The relationship is not symmetrical: When the net payoff to the signaler is marginal, a non-signaling equilibrium, in which the receiver always responds positively, is likely to occur.

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

Krebs and Dawkins (1984) predicted that signaling would be costly if a conflict of interests existed; strictly speaking the results do not support nor contradict their prediction, as no signaling occurred in the conflict-of-interest cases. It might be the case that conflicts of interest in the context of a different signaling game would indeed result in costly signals. However, the failure to evolve communication given conflicts of interest in this simple game strongly suggests that in many natural contexts (e.g., food calls, alarm calls) reliable signaling should not be expected unless it is in the interests of both parties. This conclusion is not altered by separate consideration of the specific costs of producing a signal and of making an appropriate response to that signal.

The results from the second simulation model do not confirm Krebs and Dawkins's conspiratorial whispers theory, but they definitely suggest a modification of it. As Figure 12 shows, when the net payoff to the receiver is marginal, receivers will be sceptical and express "sales-resistance" by responding only to costly signals; signalers in turn will be prepared to invest more energy in "convincing" receivers to respond positively. When communication is unambiguously good for both parties, signals are cheaper and response thresholds lower. Therefore both expensive hype and conspiratorial whispers are expected to evolve, but in a much smaller region of the payoff space than Krebs and Dawkins's theory suggests, i.e., within the cooperative region. Expensive hype is what happens when honest signaling is highly profitable to the signaler but only marginally so to the receiver. For example, suppose that a juvenile benefits by honestly signaling extreme hunger to its parent, because the parent responds by feeding it. If the net inclusive-fitness payoff to the parent is only slight, perhaps because the parent is the ostensible father and the species has a high ratio of extra-pair copulations, then costly signals by the juvenile are expected. Thus the model predicts that chicks should beg more loudly to their fathers than to their mothers, for instance.

4 VARIATIONS ON THE CONTINUOUS-SIGNAL-COST GAME

Caryl (1987) has expressed dismay at a tendency in the theoretical-biology literature for those who build mathematical or simulation models to engineer them solely in order to support a favored hypothesis, and to fail to consider the broader implications and predictions of such models. Caryl's point is that it is very easy to judiciously choose parameter values in order to get a desired result but harder to construct a model that makes sensible predictions in a range of contexts. A number of variations of the evolutionary simulation model with continuous signal and threshold values will therefore be presented. It is hoped that the results from these variants will increase the reader's confidence in the validity and general applicability of the model.

In order to avoid any further profusion of graphs, the variants will incorporate only non-signaling initial conditions. Rather than requiring the reader to constantly compare each figure with Figure 6—the mean communication index data for the continuous-signal-cost game with non-signaling initial conditions—the communication index results in each variant will be presented as *differences* from that graph. That is, Figure 6 will be used as a reference level of communication; positive results for a variant will indicate a greater relative level of communication and not an absolute measure.

4.1 Noise and uncertainty

The use of continuous values for the cost of signals and for the response threshold suggests the possibility of random noise in the signaling channel. In the real world signals will not always be accurately perceived, and Johnstone (1994) found that modeling noise or perceptual error in a signaling game in fact altered the predictions about which strategies were expected to be stable. It was thought that perhaps the inclusion of noise would alter the region of the payoff space in which communication evolved.

Noise was implemented by adding a random gaussian value ($\mu = 0$) to the energy level of the signal before it was perceived by the receiver. Thus, signals will sometimes be heard as "louder" or "softer" than they in fact are. When the random gaussian value had a standard deviation of 0.2, noise made very little difference to the communication index data, i.e., communication evolved much as in Figure 6. When the standard deviation was set to 2.0, on the other hand, communication was entirely disrupted. Presumably intermediate levels of noise would have led to a pro-

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

gressive degradation of communication, but as yet only two noise settings have been investigated. However, there was no evidence that the addition of noise could lead to honest signaling in regions of the payoff space where it would otherwise not have occurred.

Randomness was also applied to the payoff values P_S and P_R in order to investigate the effects of realistic uncertainty. The payoff values, as in all game-theoretic accounts, are intended to be average expected payoffs. However, computer simulation allows us to assign payoffs in a particular interaction that are drawn from a random gaussian distribution. Thus the long term mean will be as specified, e.g., $P_S = 2$ and $P_R = 2$, but the rewards for successful communication in any one game will be somewhat unpredictable. When the standard deviation of the random gaussian was 0.2, the evolution of stable communication was unaffected. When the standard deviation was increased to 2.0, communication started to degrade as shown in Figure 14. However, there was again no suggestion that the modeling of uncertainty in payoff values could lead to communication where it would not have otherwise evolved.

4.2 Exploitation of sensory biases and mutational lag

The simple games and simulations described here are in one sense an unfair way to test Krebs and Dawkins's (1984) conspiratorial whispers hypothesis. Krebs and Dawkins discuss the likely evolution of signals in complex real-world cases and can therefore appeal to the exploitation of response patterns that had originally been selected for other purposes, the effects of differing mutation rates in signalers and receivers, etc. Communication in their predicted costly signaling arms races was not necessarily expected to be stable. For example, in a real-world situation where it was not in the interests of receivers to respond positively to a particular signal from a predator, they might nevertheless continue to do so for some time if the signal was structurally similar to a mating signal made by members of the same species. The manipulative signaling system would break down as soon as an appropriate sequence of mutations resulted in organisms that could distinguish between the predator's signal and the conspecific mating signal. In the signaling models presented all this complexity is abstracted into the base fitness payoffs for signalers and receivers.

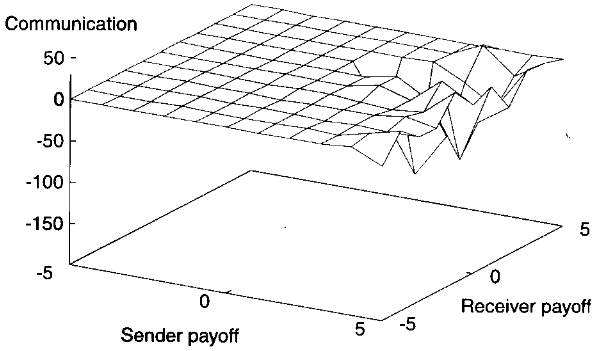


Figure 14. Difference in mean communication index between uncertain payoff variant ($\mu = 2.0$) and standard continuous-signal-cost game; non-signaling initial conditions. Each point is the difference between two means, each calculated over 25 runs. Graph rotated for clarity—cooperative quadrant appears at top right.

In an attempt to investigate these issues, two simple modifications were made to the standard continuous-signal-cost game. In the first of these, we suppose that the receivers have some other ecological reason for having a low threshold value, e.g., that the same sensory mechanisms are involved in food detection. This opens up an opportunity for signalers to exploit a "sensory bias" (Guilford & Dawkins, 1991; Ryan & Rand, 1993) in the receivers. Selection pressure for low thresholds (T) was implemented by giving receivers in each game an energy bonus (B) as follows:

$$B = 0 \text{ if } T > 5$$
$$B = 1 \text{ if } T < 0$$
$$B = (5 - T) / 5 \text{ if } 0 \leq T \leq 5.$$

The results of simulation runs of this variant are shown in Figure 15 (using Figure 6 as a baseline). When receivers have other reasons for maintaining a low response threshold, communication evolves much more reliably in the usual cooperative region of the payoff space, and also occurs in the selfish region. That is, signalers are able to manipulate receivers to their own (the signalers') advantage. Furthermore, as predicted by Krebs and Dawkins (1984), the most costly signals indeed occurred when communication had been established despite a conflict of interests.

In another variant, it is supposed that response

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

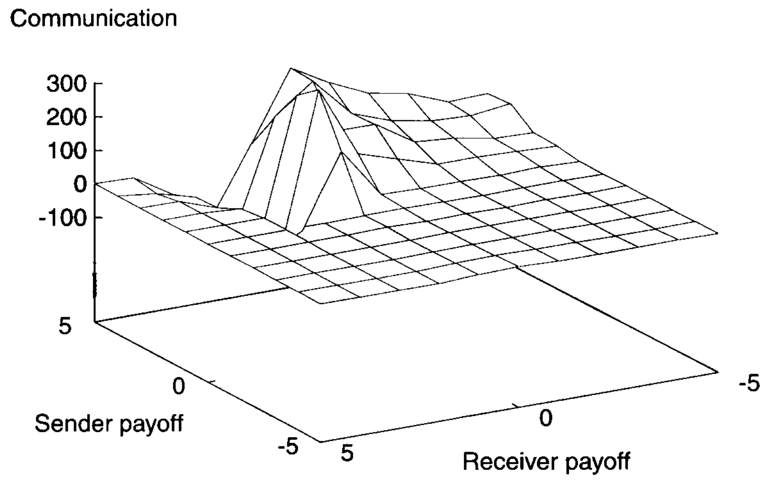


Figure 15. **Difference in mean communication index between sensory bias variant and standard continuous-signal-cost game; non-signaling initial conditions. Each point is the difference between two means, each calculated over 25 runs. Graph rotated for clarity—cooperative quadrant appears at top left.**

strategies might evolve more slowly than signaling strategies, i.e., there is a mutational lag on reception thresholds relative to signal cost values. Such a state of affairs could come about in the real world if the sensory equipment used to detect signals was older and affected by a larger network of genes than the organs used for signaling. It would then be possible that signalers might "out-evolve" receivers, and succeed in getting them to respond to selfish, manipulative signals. The idea was implemented by reducing both of the mutation rates for reception thresholds by a factor of ten. That is, the real-valued threshold gene in a newborn individual was perturbed by a random gaussian value, $\mu = 0$, $\sigma = 0.005$, and 0.1% of the time (i.e., a mutation rate of 0.001) a completely new threshold value was generated in the range ±5. The results are shown in Figure 16.

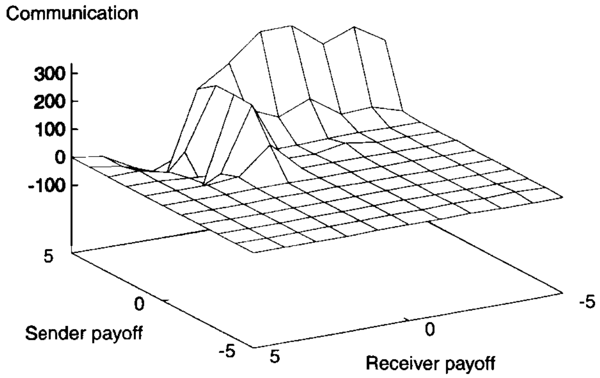


Figure 16. **Difference in mean communication index between mutational lag variant and standard continuous-signal-cost game; non-signaling initial conditions. Each point is the difference between two means, each calculated over 25 runs. Graph rotated for clarity—cooperative quadrant appears at top left.**

As with the sensory bias variant, communication is established more strongly in part of the cooperative region, but it also evolves in the selfish region for high values of P_S. Again, the most costly signals were also found when selfish communication had evolved. A puzzling feature of the result is that it does not appear to have come about simply because the low rate of mutation for threshold values meant that 500 generations was insufficient time for the optimal value to be reached. Mean threshold values when $P_S = 5$ and $P_R < 0$ were approximately four in both the mutational lag variant and the original simulation data.

4.3 The effects of spatial arrangement

Ackley and Littman (1994) and Oliphant (1996) both found that arranging signaling populations in space led to a greater degree of altruistic signaling. In Ackley and Littman's model individuals lived in small

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

groups, communicating and breeding only with their group-mates, but occasionally migrating to another nearby group. There was no spatial arrangement within each group, but the groups themselves were laid out on a grid. In Oliphant's model individuals were arranged in a ring, and were likely to communicate and to breed with their neighbors.

A spatial variant was implemented by arranging the population of 100 individuals on a toroidal 10 x 10 grid. Individuals interacted only with their eight neighbors: in each game, a signaler was chosen at random from the population and a receiver was chosen at random from among the signaler's neighbors. Breeding was also local. When one generation replaced another, the parent of the individual who would occupy a particular square was chosen, using roulette-wheel selection according to fitness, from among the nine local candidates from the previous generation. That is, the parent of the occupant of a given square would either be the previous occupant or one of the previous occupant's neighbors. The results for the spatial variant are shown in Figure 17.

Arranging the population in space leads to an increase in the reliability of communication, but only in that section of the cooperative region where honesty has already been observed to evolve. The agents have clearly not been induced to participate in altruistic communication with their neighbors. There is no communication even when signalers are merely ambivalent ($P_S = 0$). However, it can be shown that altruism of a sort has occurred. Figure 18 shows the difference in mean fitness between the spatial variant and the original simulation. There is a spike of increased fitness in the altruistic quadrant at the front of the graph: This occurs because receivers are refraining from constant positive responses, and thus being altruistic towards the signaling neighbors who would

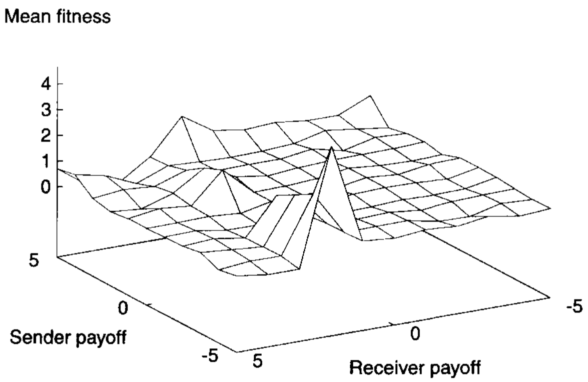


Figure 18. Difference in mean fitness between spatial variant and standard continuous-signal-cost game; non-signaling initial conditions. Each point is the difference between two means, each calculated over 25 runs. Graph rotated for clarity—cooperative quadrant appears at top left.

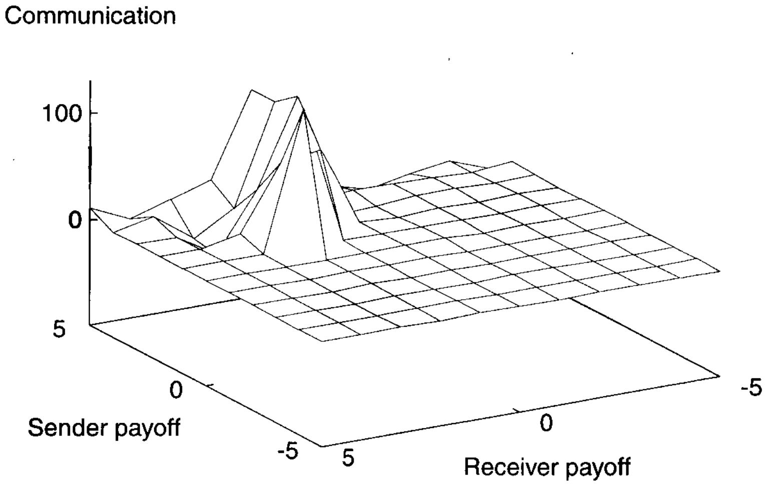


Figure 17. Difference in mean communication index between spatial variant and standard continuous-signal-cost game; non-signaling initial conditions. Each point is the difference between two means, each calculated over 25 runs. Graph rotated for clarity—cooperative quadrant appears at top left.

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

be penalized by a positive response because of the negative value of P_S in this area.

4.4 Insistent signalers

The signaling game used is not likely to be a universal model of all possible communicative interactions. In particular, and despite having the same basic structure with two signals possibly used to transmit information about a binary hidden state, the signaling game is different from those employed by Hurd (1995), Oliphant (1996), and Bullock (1997). Hurd's game, for instance, models sexual signaling, and the male signaler is not ambivalent about the female receiver's response when the hidden state is low; the signaler always prefers a positive response. A low hidden state maps to low male quality, a positive response represents a copulative episode, and even low-quality males want mating opportunities. The current signaling game, in contrast, cannot model so-called "handicap" signaling, because low-state signalers do not care about what the receiver does. Furthermore, in previous games, receivers are explicitly rewarded for accuracy in discerning the hidden state, but the game presented here allows the ecologically plausible outcome that receivers simply become disinterested in the signal. The current game is a reasonable model of situations such as alarm calls and food calls, in which potential signalers have no reason to care about what receivers do when no predator has been sighted or no food source has been found. Whereas Hurd's game serves as a (discrete) model of situations where signalers vary on some dimension, the current game models situations where signalers fall into two groups, only one of which is relevant to the potential response.

However, it is a simple matter to alter the present game such that signalers are always interested in gaining a positive response. The payoff matrix is altered such that P_S, the payoff to the signaler, is awarded whenever the receiver responds positively, regardless of the value of the hidden state. However, receivers are still only awarded their payoff, P_R, when they respond positively and the hidden state is high. There is thus a different kind of conflict of interests between the signaler and receiver.

Making signalers want positive replies all the time in this way almost completely breaks down communication (graph not shown). There are no circumstances in which receivers can trust signalers, and extreme response strategies (always responding positively or always responding negatively) are formulated purely on the basis of the payoff to the receiver. Interestingly, communication can be salvaged if the

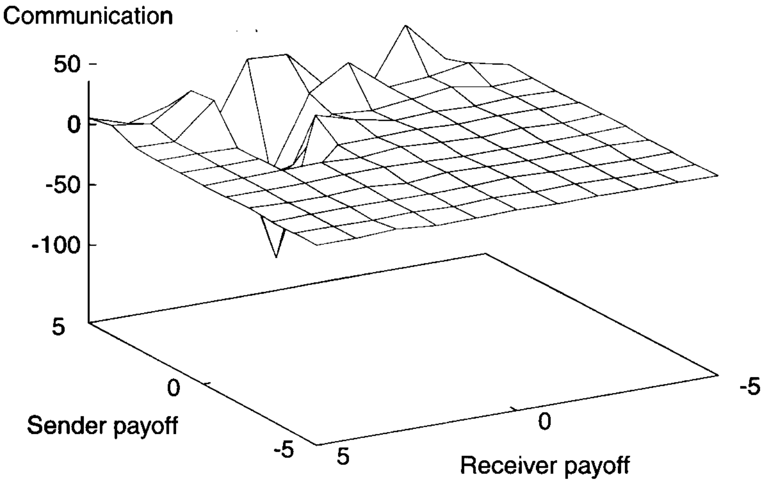


Figure 19. Difference in communication index between handicap principle variant and standard continuous-signal-cost game; non-signaling initial conditions. Each point is the difference between two means, each calculated over 25 runs. Graph rotated for clarity-cooperative quadrant appears at top left.

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

conditions of the handicap principle are applied: that is, if the unit cost of giving a signal in the low state is greater than for the high state. The results for a run in which signals in the low state cost 5 times their normal value are shown in Figure 19; relative to the standard game, communication levels are only somewhat degraded.

5 GENERAL DISCUSSION

The results from simulations of the simple and continuous-cost signaling games suggest that communication will not evolve when there is a conflict of interests between signalers and receivers. Even when signalers and receivers share a common interest, the evolution of communication is not straightforward. Firstly, receivers may fall into blindly optimistic strategies (i.e., always responding positively) that are less efficient than the communicative equilibrium but nevertheless stable. This is particularly likely to occur when the net payoff to the receiver is high. (The expected payoff for always responding positively will of course depend on the relative frequency of high and low hidden states, a factor that was not varied in the models presented). Secondly, communication may evolve but the signals involved will be more or less costly depending on the marginal payoff of the receiver, as discussed in Section 3.3.

Variations on the continuous-cost signaling game, while only briefly explored, suggest that communication can in fact evolve under conditions of conflicting interest if receivers have a sensory bias that maintains low response thresholds, or if response strategies do not evolve as quickly as signaling strategies. In these two cases, manipulative or selfish communication can occur. Of course, in the case of a sensory bias, the communication that evolves is not really occurring under a conflict of interests, because receivers are choosing the strategy that maximizes their two sources of fitness: the communication game and the independent response bias. However, an observer unaware of the receivers' response bias would observe agents responding to signals in a way that was not in their immediate interests.

Altruistic communication (considered from the point of view of signalers) was not observed under any circumstances, including the spatial variant simulation. Spatial arrangement of the population would seem not to be a guarantee of kin-selected altruism. The occurrence of apparently altruistic food and alarm calls in nature, in circumstances where reciprocal altruism and kin selection cannot be invoked, therefore remains to be explained.

The evolutionary simulation models presented are unusual in their use of non-random initial conditions. The use of non-signaling initial conditions in particular can be seen as an attempt to get at the origin or emergence of communication rather than just studying the conditions for its stability, as does orthodox game theory. Non-signaling initial conditions embody the assumption that communication must emerge from a non-communicative context—the unclamping of signaling strategies after a period of preliminary evolution can be seen as the introduction of a mutation that allows the *possibility* of signaling. The simulation results have certainly demonstrated that the conditions for stability can be very different from those for emergence.

A final qualification must be made concerning the results: The way that conflicting and congruent interests have been defined may be too simplistic. In the simple signaling game, it is true that with positive net payoffs to the signaler and the receiver, and if the hidden state is high, both agents will benefit from a positive response, and they therefore have congruent interests. However, if we consider the moment before the hidden state has been determined, it is not clear whether the interests of the two agents conflict or not. If the signaler, for example, could somehow choose the strategy of its opponent, the receiver, it would want the opponent to play an "always respond positively" strategy—that way the signaler would always receive the payoff and would not have to expend energy in signaling. However, the receiver, if similarly allowed to determine the signaler's strategy, would prefer that the signaler used an honest strategy, precisely so that the receiver could avoid the costs of responding positively to the low hidden state. Recall that Trivers (1974) defined a conflict of interests as an interaction in which natural selection favors a different outcome for each participant. It seems that the signaler and receiver in this situation favor different strategies in their opponent, and thus have a conflict of interests, even though a high value of the hidden state would mean that their interests became congruent. If this strategy-based definition of conflicting

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

interests were adopted, any situation in the cooperative payoff region, assuming signaling had a positive cost, would involve a conflict of interests—this would in turn mean that *all* of the signaling observed in the simulation models evolved despite a conflict of interests. The problem is perhaps that Trivers's (1974) and Maynard Smith and Harper's (1995) definitions are not specific enough about just what constitutes an "outcome" of the signaling game. The simpler definition of conflicting interests, as used in the body of the paper, is useful in isolating the cooperative region of payoff space as the place to expect signaling. It is not yet clear how the results should be interpreted if the strategy-based definition of conflicting interests was pursued.

NOTES

[1] As Di Paolo (1997) has pointed out, both Ackley and Littman (1994) and Oliphant (1996) do not formally demonstrate that kin selection has affected the course of evolution in their models. They simply assume that kin selection can be equated with spatial arrangement, in which an agent interacts with neighbors who are likely to be relatives.

ACKNOWLEDGEMENTS

Much of this work was completed while I was a graduate student in the School of Cognitive and Computing Sciences at the University of Sussex. I am grateful to the Association of Commonwealth Universities and the British Council for financial support at that time. I would also like to thank Seth Bullock and Ezequiel Di Paolo for valuable discussions.

REFERENCES

Ackley, D. H., & Littman, M. L. (1994). Altruism in the evolution of communication. In R. A. Brooks & P. Maes (Eds.), *Artificial Life IV*, pp. 40-48. Cambridge, MA: MIT Press.

Axelrod, R. (1984). *The Evolution of Cooperation*. New York: Basic Books.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211*, 1390-1396.

Bullock, S. (1997). An exploration of signalling behaviour by both analytic and simulation means for both discrete and continuous models. In P. Husbands & I. Harvey (Eds.), *Proceedings of the Fourth European Conference on Artificial Life (ECAL'97)*, pp. 454-463. Cambridge, MA: MIT Press / Bradford Books.

Caryl, P. G. (1987, July). Acquisition of information in contests: The gulf between theory and biology. Paper presented at the ESS Workshop on Animal Conflicts, Sheffield, UK.

Dawkins, R. (1976). *The Selfish Gene*. Oxford: Oxford University Press.

Di Paolo, E. A. (1997). An investigation into the evolution of communication. *Adaptive Behavior, 6*(2), 285-324.

Grafen, A. (1990). Biological signals as handicaps. *Journal of Theoretical Biology, 144*, 517-546.

Guilford, T., & Dawkins, M. S. (1991). Receiver psychology and the evolution of animal signals. *Animal Behavior, 42*, 1-14.

Hamilton, W. D. (1964). The genetical evolution of social behavior. *Journal of Theoretical Biology, 7*, 1-52.

Hinde, R. A. (1981). Animal signals: Ethological and games-theory approaches are not incompatible. *Animal Behavior, 29*, 535-542.

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

Iwasa, Y., Pomiankowski, A., & Nee, S. (1991). The evolution of costly mate preferences II. The "handicap" principle. *Evolution, 45*(6), 1431-1442.

Johnstone, R. A. (1994). Honest signaling, perceptual error and the evolution of 'all-or-nothing' displays. *Proceedings of the Royal Society of London: Biological Sciences, 256*, 169-175.

Krebs, J. R., & Dawkins, R. (1984). Animal signals: Mind reading and manipulation. In J. R. Krebs & N. B. Davies (Eds.), *Behavioral Ecology: An Evolutionary Approach* (Second edition), pp. 380-402. Oxford: Blackwell.

MacLennan, B. J., & Burghardt, G. M. (1994). Synthetic ethology and the evolution of cooperative communication. *Adaptive Behavior, 2*(2), 161-188.

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

Maynard Smith, J. (1993). *The Theory of Evolution* (Canto edition). Cambridge: Cambridge University Press.

Maynard Smith, J., & Harper, D. G. C. (1995). Animal signals: Models and terminology. *Journal of Theoretical Biology, 177*, 305-311.

Oliphant, M. (1996). The dilemma of Saussurean communication. *BioSystems, 37*, 31-38.

Ryan, M. J., & Rand, A. S. (1993). Sexual selection and signal evolution: The ghost of biases past. *Philosophical Transactions of the Royal Society of London: Biological Sciences, 340*, 187-195.

Seyfarth, R., Cheney, D. L., & Marler, P. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science, 210*, 801-803.

Sherman, P. W. (1977). Nepotism and the evolution of alarm calls. *Science, 197*, 1246-1253.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*, 35-57.

Trivers, R. L. (1974). Parent-offspring conflict. *American Zoologist, 14*, 249-264.

Werner, G. M., & Dyer, M. G. (1992). Evolution of communication in artificial organisms. In C. G. Langton, C. Taylor, J. D. Farmer & S. Rasmussen (Eds.), *Artificial Life II*, pp. 659-687. Redwood City, CA: Addison-Wesley.

Wheeler, M., & de Bourcier, P. (1995). How not to murder your neighbor: Using synthetic behavioral ecology to study aggressive signaling. *Adaptive Behavior, 3*(3), 273-309.

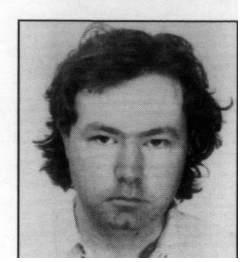
Williams, G. C. (1966). *Adaptation and Natural Selection*. Princeton, NJ: Princeton University Press.

Zahavi, A. (1975). Mate selection—A selection for a handicap. *Journal of Theoretical Biology, 53*, 205-214.

Zahavi, A. (1987). The theory of signal selection and some of its implications. In V. P. Delfino (Ed.), *International Symposium on Biological Evolution*, pp. 305-327.

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.

ABOUT THE AUTHOR



Jason Noble

Jason Noble gained a B.A. (Hons.) in psychology from Macquarie University, Sydney, in 1992. After pursuing independent musical interests for several years, he was awarded a Commonwealth Scholarship in 1995, and joined the doctoral program in the School of Cognitive and Computing Sciences at the University of Sussex. He was interested in using evolutionary simulation techniques to look at the origins of animal communication systems, and completed his thesis in 1998. He now holds a postdoctoral position with the Center for Adaptive Behavior and Cognition, part of the Max Planck Institute for Human Development in Berlin. His current research is on the evolution of simple cognitive mechanisms in social information gathering.

Downloaded from http://adb.sagepub.com at Periodicals Office on February 29, 2008
© 1999 International Society of Adaptive Behavior. All rights reserved. Not for commercial use or unauthorized distribution.