Research

CrossMark
click for updates

Author for correspondence:
Kevin J. S. Zollman
e-mail: kzollman@andrew.cmu.edu

Royal Society **Publishing**
Informing the science of the future

Between cheap and costly signals: the evolution of partially honest communication

Kevin J. S. Zollman[1], Carl T. Bergstrom[2,3] and Simon M. Huttegger[4]

[1]Department of Philosophy, Carnegie Mellon University, Pittsburgh, PA 15213-3890, USA
[2]Department of Biology, University of Washington, Seattle, WA 98195-1800, USA
[3]Sante Fe Institute, Santa Fe, NM 87501, USA
[4]Department of Logic and Philosophy of Science, University of California, Irvine, CA 92697, USA

Costly signalling theory has become a common explanation for honest communication when interests conflict. In this paper, we provide an alternative explanation for partially honest communication that does not require significant signal costs. We show that this alternative is at least as plausible as traditional costly signalling, and we suggest a number of experiments that might be used to distinguish the two theories.

1. Introduction

Communication is ubiquitous in the biological world. When the interests of signaller and signal receiver are perfectly aligned, the evolutionary benefits of reliable communication are straightforward. But when interests are not aligned, signallers might be selected to manipulate signal receivers with misleading signals, and the signal receivers might evolve to disregard such communications. Why does communication not break down in situations that involve conflict?

Costly signalling theory provides one explanation. Communication is framed as a signalling game; with appropriate signal costs, honest communication is a Nash equilibrium of this game. If signalling is sufficiently costly so that lying is too costly to be worthwhile, but honest signals are not too costly to send, all signallers may choose to signal honestly at equilibrium [1–3].

However, the costly signalling explanation for honest communication is problematic for several reasons. On the theoretical side, costly signalling can be a very expensive mode of information exchange, and at equilibrium signals can be so costly that all involved would be better off simply not communicating [4]. From a dynamical perspective, the necessary level of cost may make it difficult to evolve signalling at all [5]. On the empirical side, researchers have not always been able to find substantive signal costs associated with putative costly signal systems—despite evidence that these systems do convey, at least, some information among individuals with conflicting interests [6–14]. What, then, are we to make of empirical situations in which signals appear to be informative even without the high costs required by costly signalling models?

One approach to resolving this problem is to recognize that—contra early work on the problem—the costly signalling mechanism does not require a cost to the signals that are actually sent in equilibrium. It requires only that out-of-equilibrium signals, i.e. 'lies', be too costly to be worthwhile [15,16]. In other words, it is not the cost of signalling, but rather the marginal cost of signalling, that ensures honesty [17].

In this paper, we offer an alternative explanation. We show that the paradigmatic costly signalling games from the animal behaviour literature allow a 'hybrid' signalling equilibrium, in which inexpensive signals are able to support a partially informative signalling equilibrium. While the existence of hybrid equilibria in economic models of signalling has been known for some time [18,19], the evolutionary significance of these equilibria is only just now being considered. Initial analysis focused on the Sir Philip Sidney game [5] and a signalling game from economics known as the Spence signalling game [20]. In this study, we
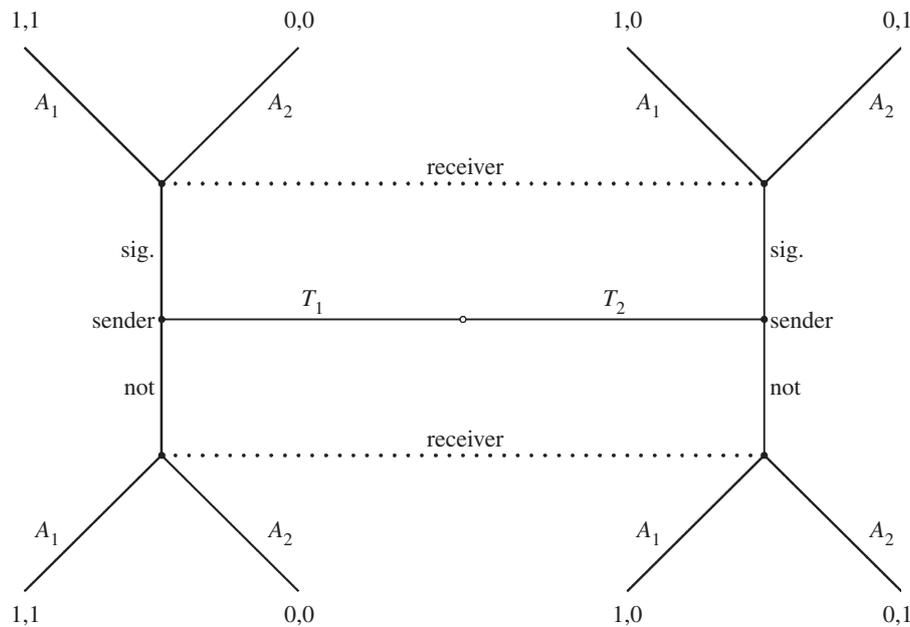
Figure 1. An action–response game with cost-free signals and a partial conflict of interest. The game begins at the central node (open circle). The first move is a move by 'nature' to determine the type of the signaller; this type is revealed to the signaller but not the receiver. In the second move, the signaller conditions its behaviour on its type and chooses whether or not to send a signal. As the third move, the receiver must choose between two actions. The receiver can condition on the signal, but not the type; this uncertainty is represented by the dotted lines.

demonstrate that hybrid equilibria in biological settings do not depend on the particular relatedness structure of the Sir Philip Sidney game, but rather are of general importance in signalling games of many types. We show that hybrid equilibria exist both in *differential cost models* such as models of mate advertisement [2], and in *differential benefit models* such as models of offspring begging [3,21]. Moreover, we demonstrate that under reasonable evolutionary dynamics, the hybrid equilibrium has a basin of attraction comparable in size with that of the classic costly signalling equilibrium. We conclude with a discussion of the implications of this work for empirical studies. We argue that most empirical studies taken as evidence of costly signalling theory fail to discriminate between the classical costly signalling equilibrium and the hybrid equilibrium presented here. However, these two different types of signalling equilibria make different predictions regarding equilibrium behaviour. This paves the way for future studies that may be able to empirically distinguish among these hypotheses.

2. Low-cost signals in costly signalling games

(a) Differential costs and differential benefits

Figure 1 illustrates a canonical two-player signalling game in extended form. Player 1, the *sender* is one of two types, T_1 or T_2. In the case of signalling between relatives, as exemplified by Maynard Smith's Sir Philip Sidney game, the types might represent the state of need of the receiver. In signalling between mates or signalling between predators and prey, these types might represent the quality of the signaller. Although the sender can condition her behaviour on her type, player 2, the *receiver*, cannot directly observe the signaller's type. Instead, the receiver can observe only whether a signal, designated *Sig*, is produced or not. On the basis of this observation, the receiver can choose one of two actions, A_1 or A_2. In the Sir Philip Sidney game, the action A_1 can be interpreted as transferring a resource and A_2 can be interpreted

Table 1. All possible pure strategies in the action–response game pictured in figure 1.

label	description
S_1	signal if T_1 and do not signal if T_2
S_2	signal always
S_3	never signal
S_4	signal if T_2 and do not signal if T_1
R_1	A_1 if signal is observed, A_2 otherwise
R_2	A_2 always
R_3	A_1 always
R_4	A_2 if signal is observed, A_1 otherwise

as not transferring. In mate selection, A_1 represents mating, whereas A_2 represents not mating. In predator–prey signalling, A_1 denotes declining pursuit and A_2 denotes pursuit. This game features four pure (conditional) strategies for the sender, and four for the receiver. These are listed in table 1.

This game structure reveals that there is a common core to signalling among relatives, signalling to potential mates or signalling to predators. The pay-offs of this game result in a partial conflict of interest between signaller and receiver. When the signaller is type T_1 (needy, a high-quality mate, or uncatchable as prey), both the sender and receiver prefer the same actions: both players prefer that the receiver take action A_1. However, when the sender is of type T_2, the sender and the receiver have divergent interests. In this case, the sender still prefers that the receiver take action A_1, but a fully informed receiver would prefer to take action A_2. This partial conflict of interest introduces the evolutionary conundrum: a T_1 type individual benefits by advertising its true quality to the receiver, whereas a T_2 type individual benefits by deceiving the receiver. The game illustrated in
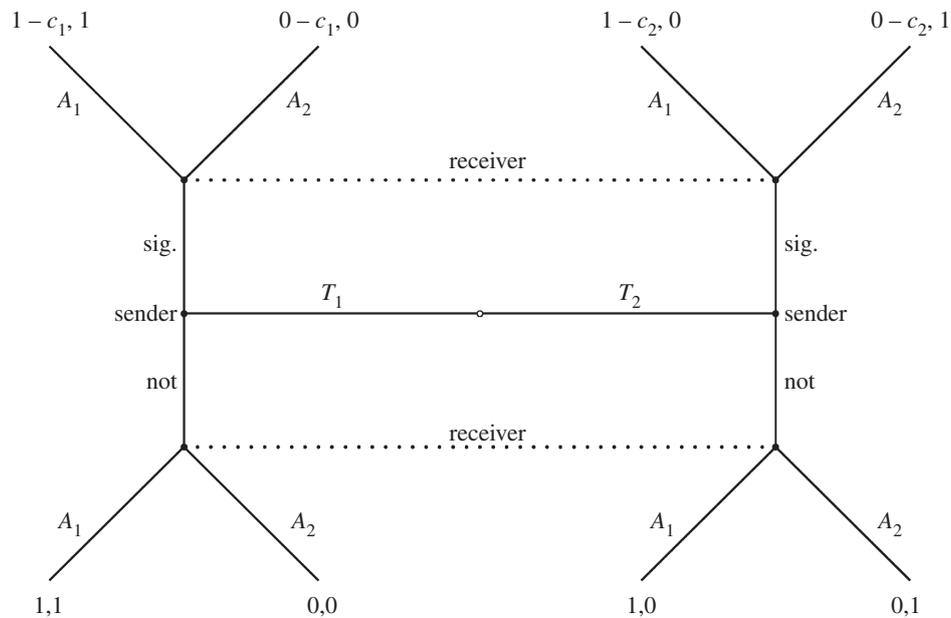
Figure 2. A differential costs signalling game. Here, the game from figure 1 is modified by adding condition-dependent signal costs to the sender's pay-offs, where c_1 represents the cost for senders of type T_1 to send the signal, and c_2 represents the cost to senders of type T_2.

Table 2. The normal form of the differential costs signalling game in figure 2. The strategy labels are defined in table 1.

	R_1	R_2	R_3	R_4
S_1	$(x - xc_1)$, 1	$-xc_1$, $(1 - x)$	$(1 - xc_1)$, x	$-xc_1 + (1 - x)$, 0
S_2	$(1 - xc_1 - (1-x)c_2)$, x	$-xc_1 - (1-x)c_2$, $(1 - x)$	$1 - xc_1 - (1-x)c_2$, x	$-xc_1 - (1-x)c_2$, $(1 - x)$
S_3	0, $(1 - x)$	0, $(1 - x)$	1, x	1, x
S_4	$(1 - x)(1 - c_2)$, 0	$(1-x)c_2$, $(1 - x)$	$1 - (1-x)c_2$, x	$x - (1-x)c_2$, 1

figure 1 has no cost associated with signalling. Without signal costs, this conflict cannot be resolved by honest communication. There is no stable set of strategies where the type is revealed by the signal. All the pure-strategy equilibria of this game are known as pooling equilibria, where both types send the same signal.

The central insight of costly signalling theory is that signal costs can stabilize honest communication. For this to work, different types of signallers must have different incentives to send signals. Such differing incentives can arise as a result of differential costs, or differential benefits, or both. In *differential cost* models, one type pays lower signalling costs than the other. This is the scenario originally described by Zahavi [1] in the context of signalling between mates; Grafen [2] treated this case in his landmark 1990 paper. Signalling from prey to predator also typically involves differential costs [22,23]. In *differential benefit* models, one type reaps a larger reward from the receiver's response than the other. Differential benefits drive honest signalling in Maynard Smith's Sir Philip Sidney game [3] and in Godfray's model of nestling begging [24]. In what follows, we treat differential cost and differential benefit models in turn.

(b) Hybrid equilibria for differential cost games

Consider the game pictured in figure 2 and table 2. This game allows for the possibility of stable advertisement of type, because the signal imposes differential costs on senders of different types. Take the strategy where T_1 signals and T_2 does not. In such a case, it is obviously best for the receiver to take action A_1 when the signal is observed and action A_2 if not. So long as the T_1 individuals pay a signal cost of $c_1 < 1$, it is strictly in their interest to signal in order to ensure that the receiver takes action A_1. So long as the T_2 individuals would pay a cost $c_2 > 1$ if they were to signal, then it is strictly in their interest to refrain from signalling, because the cost of the signal outweighs the benefit obtained by inducing the receiver to take action A_1:

$$c_2 > 1 > c_1. \tag{2.1}$$

This is the familiar separating equilibrium in a differential costs model.

However, there is another equilibrium that allows for partial transfer of information. Consider the following mixed strategy pair: the signaller mixes between strategy S_1 with probability $(1 - \alpha)$ and strategy S_2 with probability α. The receiver mixes between strategy R_1 with probability β and strategy R_2 with probability $(1 - \beta)$. The resulting behaviour is illustrated in figure 3. T_1 individuals always send the signal, but T_2 individuals play a mixed strategy, sending the signal with probability $\alpha < 1$ and refraining from sending the signal with probability $(1 - \alpha)$. If the sender does not observe the signal, she takes action A_2 with certainty. If she does observe the signal, she plays a mixed strategy, taking action A_1 with probability $\beta < 1$ and taking action A_2 with probability $(1 - \beta)$.
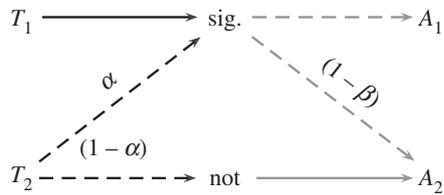
Figure 3. An illustration of the hybrid equilibrium. Black lines denote the strategy of the sender, and grey lines represent the strategy of the receiver. Solid lines represent conditionally pure strategies; dashed lines represent conditionally mixed strategies.

In plain English, this means that the sender sometimes 'lies' and is honest at other times, whereas the receiver only sometimes chooses the sender's favoured action.

When the players behave in this way, information transfer is imperfect, unlike the separating equilibrium; the signal does not indicate the type of the receiver with certainty. However, *the signaller does transmit some information to the receiver and the receiver does make some use of that information*: because T_1 signals always and T_2 signals only occasionally, observing a signal increases the likelihood that the signaller is a T_1-type. If the prior probability of T_1 is x, then the probability that a signalling individual is of type T_1 equals $x/(x+\alpha(1-x))>x$.

When does this pair of mixed strategies constitute an equilibrium? First, consider the actions of the receiver. Because only T_2 individuals fail to signal, receiving no signal informs the receiver that the signaller is a T_2 individual and that the receiver does best to take action A_2. If at equilibrium, the receiver randomizes between A_1 and A_2 in response to receiving a signal, these two actions must yield the same expected pay-off. Supposing the sender's strategy conforms to the description just offered, the expected pay-off to the receiver of playing A_1 given that a signal is observed is

$$\frac{x}{x+\alpha(1-x)}.$$

The expected pay-off of playing A_2 given that a signal is observed is

$$1 - \frac{x}{x+\alpha(1-x)}.$$

These expected pay-offs are equal when the signaller's mixed strategy uses strategy S_2 with frequency α given by $\alpha = x/(1-x)$. This condition can be satisfied only when T_1-types are sufficiently rare. (The fact that $x < \frac{1}{2}$ in this model should not be taken as some deep fact about the possibility of hybrid equilibria. Rather, the threshold of $\frac{1}{2}$ arises because we have chosen a symmetric pay-off for the receiver's success when the sender is type T_1 or type T_2. Had we made these pay-offs different, there would be a different constraint on x.)

Now consider the sender's strategy. T_1 individuals always signal. For this to be an equilibrium behaviour, the pay-off from signalling must be at least as large as that from not signalling: $\beta(1-c_1)+(1-\beta)(-c_1) \geq 0$, which obtains when $\beta \geq c_1$. Individuals of type T_2 mix between signalling and not signalling. Thus, these two behaviours must yield the same pay-off: $\beta(1-c_2)+(1-\beta)(-c_2) = 0$, which reduces to $\beta = c_2$. Given these requirements, β can take on a value in $(0,1)$ and thus represent a mixed strategy best response
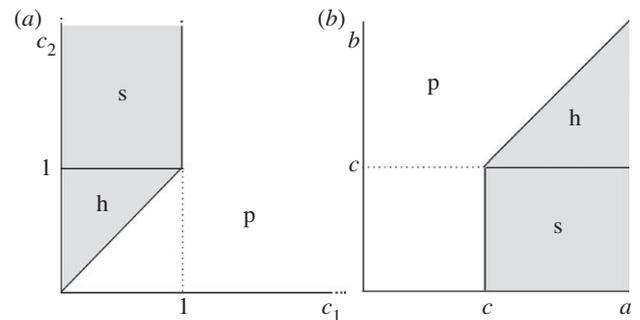

Figure 4. An illustration of the locations of separating, hybrid and pooling equilibrium for (*a*) the differential cost game and (*b*) the differential benefit game. The regions marked by s and h are regions where the separating and hybrid equilibria (respectively) exist. Pooling equilibria exist for all parameter values, but in the regions marked with p, they are the only equilibria.

when $1 > c_2 > 0$ and when $c_2 \geq c_1$. Thus, the overall conditions for the hybrid equilibrium are

$$\alpha = \frac{x}{1-x}, \tag{2.2}$$
$$1 > \beta = c_2 > 0 \tag{2.3}$$
and $c_2 \geq c_1 \tag{2.4}$$

Comparing conditions (2.1) and (2.4), we see that both the separating equilibrium and the hybrid equilibrium require $c_2 \geq c_1$. Given condition (2.3), the hybrid equilibrium exists precisely when the cost c_2 is not sufficiently high to maintain the separating equilibrium, i.e. $c_2 < 1$. Differential cost is still necessary for the hybrid equilibrium, but the signal can be considerably cheaper than in the separating equilibrium. Figure 4*a* illustrates the regions of parameter space where the hybrid and separating equilibrium exist.

How do the two players fare at the different equilibria? At the traditional costly signalling equilibrium, the receiver gets a pay-off of 1, regardless of the state of the sender. At the hybrid equilibrium, this pay-off is less, because occasionally the receiver performs action A_2 when the sender is of type T_1 and sometimes the receiver performs A_1 when the sender is of type T_2. From the perspective of the receiver, the traditional signalling equilibrium yields a higher pay-off.

Things are more complicated for the sender. Use H and S superscripts to the values c_1 and c_2 to refer to costs at the hybrid or separating equilibrium, respectively. Prior to learning her type, the sender's expected fitness in the traditional costly signalling equilibrium is $x(1 - c_1^S)$. In the hybrid equilibrium, it is $x(c_2^H - c_1^H)$. Whether it is better for the sender to be in a signalling or hybrid equilibrium depends on *which* signalling and hybrid equilibrium we compare. This much can be said: for any signalling equilibrium, there exist hybrid equilibria where the sender fares better than in that signalling equilibrium.

(c) Hybrid equilibrium for differential benefit games

Differential costs are not essential for a costly signalling equilibrium. Stable signalling can also arise when the two types of signallers face the same signal costs, but accrue different benefits from the receiver's response. This is the scenario modelled in Maynard Smith's Sir Philip Sidney game [3]. A differential benefits signalling game is illustrated in figure 5 and table 3.
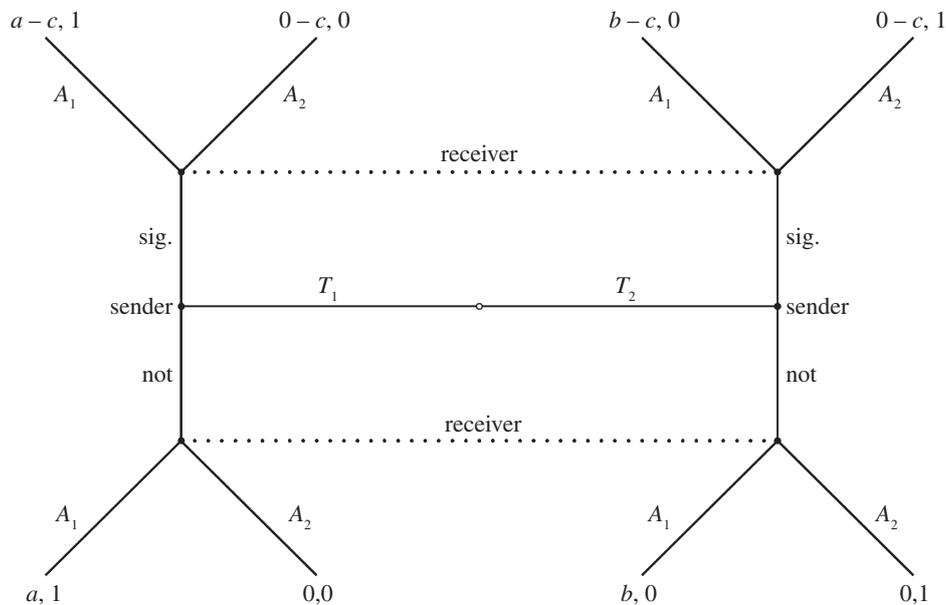
Figure 5. A differential benefits signalling game. The values a and b represent the benefit of taking action A_1 to T_1 and T_2 receivers, respectively. Irrespective of the signaller's type, the cost of signalling is c.

In this differential benefits model, the two types pay the same cost c of signalling, but the benefit they receive from getting a receiver to take action A_1 varies. The separating equilibrium requires first that the T_1 individuals do best by signalling, namely that $a > c$. Second, it requires that the T_2-types do better by not signalling, i.e. that $c > b$. Thus, the benefit of A_1 must be higher for the T_1 individuals than it is for the T_2 individuals, and the cost of signalling must separate the two:

$$a > c > b. \tag{2.5}$$

There is also a hybrid equilibrium. Consider the same combination illustrated in figure 3: T_1 individuals always signal, whereas T_2 individuals mix between signalling and not signalling; receivers mix between A_1 and A_2 in response to a signal, and always select A_2 in the absence of a signal. The condition on α remains the same as in the differential costs model: $\alpha = x/(1 - x)$.

The condition on the benefits is as follows: $\beta = c/b \geq c/a$. The conditions are satisfied with β taking a value in $(0,1)$ when $a \geq b > c$. The overall conditions for a hybrid equilibrium are thus

$$\alpha = \frac{x}{1-x}, \tag{2.6}$$

$$a \geq b > c \tag{2.7}$$

and $\beta = \frac{c}{b}$. $\tag{2.8}$

Comparing conditions (2.5) and (2.7), we see that both the separating and hybrid equilibria require $a \geq b$, i.e. that the benefit from A_1 be greater for T_1 individuals than for T_2 individuals. However, the hybrid equilibrium exists precisely when the signal cost c is too low to sustain the separating equilibrium.[1] Figure 4b illustrates the regions of parameter space in which each of these equilibria are found for differential cost models.

Again, we can compare the pay-offs obtained under the hybrid and signalling equilibria. As in the differential cost model, the receiver invariably gets a higher pay-off at the separating equilibrium. In the differential benefit model, the expected fitness of the sender is $x(a - c^S)$ at the separating equilibrium and $xc^H(a/b - 1)$ at the hybrid equilibrium. As with the differential cost model, for every separating equilibrium, there exist hybrid equilibria at which the sender fares better. In both models, separating can be so costly that signallers do better with reduced information transfer. Because in this game—unlike in the Sir Philip Sidney game—the signal cost does not decrease the pay-off to the receiver, the receiver is always better off receiving more information.

Many biologically relevant interactions will feature both differential cost and differential benefit. Although we have divided this into two cases for the purposes of clarity, the same hybrid equilibrium exists in a model that combines both the differential cost and differential benefit. The conditions for its existence are more complex but essentially the same conditions as above.

3. Evolutionary implications

The mere existence of Nash equilibria does not guarantee their evolutionary significance. As is well known, many Nash equilibria are unlikely evolutionary outcomes in that they are unstable in one sense or another [25]. In this section, we will demonstrate the evolutionary significance of the hybrid equilibria characterized earlier by illustrating that these equilibria attract open sets of nearby states in the two-population replicator dynamics.

The two-population replicator dynamics captures the basic process of evolution by natural selection in asymmetric games [25]. Let x_i be the relative frequency of sender type i and y_j be the relative frequency of receiver type j, $i,j = 1, \ldots, 4$. This dynamics is given by:

$$\dot{x}_i = x_i(\pi_i(\mathbf{y}) - \pi(\mathbf{x}, \mathbf{y})), \tag{3.1a}$$

and

$$\dot{y}_j = y_j(\pi_j(\mathbf{x}) - \pi(\mathbf{y}, \mathbf{x})). \tag{3.1b}$$

Table 3. The normal form of the differential benefits signalling game in figure 5. The strategy labels are defined in table 1.

	R_1	R_2	R_3	R_4
S_1	$x(a - c),1$	$-xc,(1 - x)$	$x(a - c) + (1 - x)b,x$	$- cx + (1 - x)b,0$
S_2	$xa + (1 - x)b - c,x$	$-c,(1 - x)$	$xa + (1 - x)b - c,x$	$-c,(1 - x)$
S_3	$0,(1 - x)$	$0,(1 - x)$	$xa + (1 - x)b,x$	$xa + (1 - x)b,x$
S_4	$(1 - x)(b - c),0$	$-(1 - x)c,(1 - x)$	$xa + (1 - x)(b - c),x$	$xa - (1 - x)c,1$

The vectors $\mathbf{x} = (x_1, \ldots, x_4)$ and $\mathbf{y} = (y_1, \ldots, y_4)$ are the vectors of strategy frequencies for signaller and receiver, respectively. The expression $\pi_i(\mathbf{y})$ gives the pay-off of i against $\mathbf{y}$ and $\pi_j(\mathbf{x})$ gives the pay-off of j against $\mathbf{x}$. The average pay-off in the sender population is given by $\pi(\mathbf{x}, \mathbf{y})$ and the average pay-off in the receiver population by $\pi(\mathbf{y}, \mathbf{x})$. The two-population replicator dynamics is thus a simple formalization of the idea that, in each population, strategies with above-average fitness thrive, whereas strategies with below-average fitness decline.

We can now consider the evolutionary properties of both the signalling and hybrid equilibria. In both the differential cost game and the differential benefits game, the signalling equilibrium is a strict Nash equilibrium (when the listed strict inequalities are satisfied). Therefore, the signalling equilibrium is necessarily an attracting state in the two-population replicator dynamics [25]. Although this does not guarantee that it is a global attractor—meaning that we cannot be assured that the signalling equilibrium be reached from any ancestral state—we can conclude that a population at a signalling equilibrium will remain there if subject to only small perturbations.

The hybrid equilibria in both games are slightly more complex. On the plane defined by the four strategies (two sender and two receiver) that make up the hybrid equilibria, populations cycle around the hybrid equilibrium indefinitely. States near, but not on this plane, converge to the plane. So, the hybrid equilibrium characterizes a set of states that are also evolutionarily significant. In §3a,b, we demonstrate this for both games.

(a) Differential cost games

There are two major claims that must be established. First, we will characterize trajectories on the plane composed by strategies S_1, S_2, R_1 and R_2. Second, we will show that population states near the plane will converge to it.

We will begin by considering the game comprised by the four strategies listed earlier. This game produces the following two pay-off matrices:

$$S = \begin{pmatrix} x(1 - c_1) & -xc_1 \\ x(1 - c_1) + (1 - x)(1 - c_2) & -xc_1 - (1 - x)c_2 \end{pmatrix}$$

$$R = \begin{pmatrix} 1 & x \\ (1 - x) & (1 - x) \end{pmatrix}.$$

By subtracting a (different) constant from each column, one obtains the following equivalent game:[2]

$$S' = \begin{pmatrix} 0 & (1 - x)c_2 \\ (1 - x)(1 - c_2) & 0 \end{pmatrix}$$

$$R' = \begin{pmatrix} 0 & 2x - 1 \\ -x & 0 \end{pmatrix}.$$

The hybrid equilibrium corresponds exactly to the interior Nash equilibrium of this game (whenever it exists). The interior Nash equilibrium is either unstable for the replicator dynamics restricted to the plane, or there are closed orbits around it. Hofbauer & Sigmund [25] show that there are three necessary conditions that are required for the two-population replicator dynamics to produce the second case, i.e. cyclic behaviour. First, the non-zero entries in S' must be of the same sign. Second, the non-zero entries in R' must also be of the sign. Finally, the non-zero entries in S' and R' must be of different signs from one another. The first condition holds on assumption that $c_2 < 1$, which is required for the existence of the hybrid equilibrium. The second condition holds on condition that $x < 1/2$, again a requirement for the hybrid equilibrium. From these two requirements, the final condition follows. It is shown in Hofbauer & Sigmund [25] that when these conditions are satisfied, there exists a constant of motion. It follows that the replicator dynamics cycle around the hybrid equilibrium in closed orbits on the plane comprising the four strategies S_1, S_2, R_1 and R_2.

What about states off this plane? Consider perturbations from the hybrid equilibrium into the interior of the space. At the hybrid equilibrium, the average pay-off for the sender is $x(c_2 - c_1)$ and $(1 - x)$ for the receiver. The pay-off of S_3 and S_4 against the hybrid equilibrium is 0 in both cases; as a result, the motion in the interior close to the hybrid equilibrium will be in the direction of the hybrid equilibrium. The pay-off for strategies R_3 and R_4 against the hybrid equilibrium is x. Because the hybrid equilibrium exists only when $x < \frac{1}{2}$, it follows that whenever the hybrid equilibrium exists, the motion of the system will be towards the hybrid equilibrium from the interior when the system is close to the hybrid equilibrium. These considerations show more formally that the so-called transversal eigenvalues of the hybrid equilibrium [25] are negative. Together with the results regarding the plane comprised of S_1, S_2, R_1 and R_2, it follows that interior trajectories approach the hybrid equilibrium in a spiralling manner. The hybrid equilibrium is 'strongly stable' relative to outside states and 'weakly stable' towards states on the plane.[3]

That stable cycles are observed on the face of the simplex is structurally unstable—small changes in the underlying dynamics (as might occur with the introduction of mutation or other stochastic effects) are likely to alter this prediction of our model. A number of different outcomes can occur depending on what changes are made to the underlying dynamical system [25]. We leave the details of these changes for future research. On the other hand, the stability of the plane itself against invasion from types off the plane is unlikely to be affected by small changes in the dynamics. While we are unsure whether cycles should be expected in natural
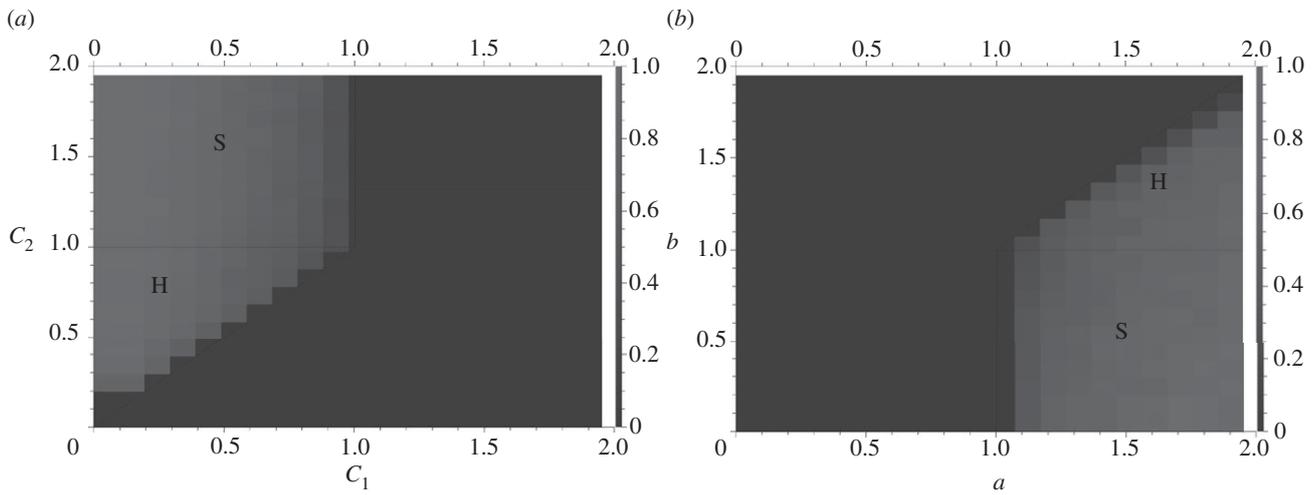
Figure 6. Simulation results establishing the basins of attraction for signalling and hybrid equilibria. In both cases, x (the probability of being T_1) is equal to 0.25. In (*b*), $c = 1$. Each point in the graph represents one setting of the variables, and the colour represents the estimated size of the basin of attraction for the replicator dynamics. (*a*) Varying cost signalling and hybrid; (*b*) varying benefit signalling and hybrid.

populations, the stability of populations on the hybrid equilibrium plane to invasion by other types suggests that one should expect to find natural populations on or near this plane.

(b) Differential benefit games

We will now complete the same analysis for the differential benefit version of the signalling game. R' remains as before, but S' is different:

$$S' = \begin{pmatrix} 0 & (1-x)c \\ (1-x)(b-c) & 0 \end{pmatrix}.$$

Again, the three conditions are satisfied when the hybrid equilibrium exists, so we can conclude that the motion on the plane is cyclic. In the differential benefit game, the average pay-off for sender in the hybrid equilibrium is $xc(a/b - 1)$, which is positive so long as $a > b$. The pay-off for S_3 and S_4 are both 0, and so we can conclude that the motion in these directions is towards the hybrid equilibrium. The analysis for the receiver is identical as above.

(c) Basins of attraction

Thus far, we have shown that a non-zero fraction of the state space will converge to the hybrid equilibrium when it exists. A similar result holds for signalling equilibria, provided that they exist. Figure 6 illustrates simulation results that capture the relative sizes of the basins of attraction for both signalling and hybrid equilibria. This graph reveals that in both the differential cost and differential benefit models, the hybrid equilibrium is as significant an evolutionary outcome as the signalling equilibrium.

In these graphs, the basin of attraction of the hybrid equilibrium is modelled by measuring how many initial states converge to the surface which only consists of strategies S_1, S_2, R_1 and R_2. We have already illustrated that once on this face, populations will enter stable cycles around the hybrid equilibrium. This represents populations that are heterogeneous with respect to their willingness to signal honestly and respond to the signal.

Note that the analysis in this section deals with the situation in which mixed strategies are manifested by population-level polymorphism of pure strategists rather than a monomorphic population of pure strategists [26]. There is no straightforward way to characterize the dynamic stability of a hybrid equilibrium manifested as a monomorphic population of mixed strategists. The hybrid equilibrium is not a strict Nash equilibrium because mixed Nash equilibria are never strict. Nor is it an evolutionarily stable strategy (ESS), because by Selten's theorem [27], there are no ESSs in mixed strategies in role-asymmetric games.

4. Discussion

We have shown that hybrid equilibria are general features of the costly signalling games studied in evolutionary biology. They arise and facilitate low-cost, low-fidelity communication not only in differential benefit games used to model nestling begging, but also in the differential cost games used to model sexual signalling and prey-to-predator communication. Our findings suggest an important empirical avenue for investigation. Experiments that purportedly demonstrate costly signalling in animal populations have traditionally compared two hypotheses: the null hypothesis—that no communication is taking place—against an alternative hypothesis—that some communication is taking place. While these studies have conclusively rejected the hypothesis that no signalling is taking place, they were not designed to distinguish between the traditional signalling hypothesis and the hybrid equilibrium hypothesis presented here.

Empirical studies could distinguish between the hybrid equilibrium and the separating equilibrium. First, the hybrid equilibrium hypothesis predicts that signalling will be honest some, but not all of the time. Moreover, the deviation from honesty should take on a particular direction that distinguishes the hybrid equilibrium from a separating equilibrium muddied by noise or mistakes. At the hybrid equilibrium, 'low' type individuals (T_2 in our model) should sometimes deviate and send the costly signal, whereas 'high' type individuals (T_1) should never deviate by failing to signal. Similarly, receivers should sometimes deviate and fail to respond positively

to a signal, but they should never deviate by responding positively in the absence of a signal.

Second, our model predicts that the population would be heterogeneous with respect to both the sending and receiving strategies. Third, our model suggests that it is unlikely that the population would be in equilibrium, but rather would be observed cycling around the equilibrium. This last prediction is perhaps the most tenuous, because the cycling behaviour observed here will not be stable to small perturbations in the underlying dynamics.

Beyond the empirical significance of our results, we have demonstrated an alternative evolutionary explanation for (partially) honest communication in situations of conflict of interest, which can range from parent–offspring interactions to mating advertisement to predator–prey interactions. Although this equilibrium had been previously observed [5,20], it was not known if this phenomena was an artefact of the particular games or if it represented a general phenomenon common to many different signalling interactions. In this paper, we show that this type of signalling is at least as evolutionarily plausible as that offered by the traditional costly signalling models and may fit better with the observed data on signal costs. In this respect, it may represent a superior theory to traditional handicap theory.

Endnotes

[1]Huttegger & Zollman [5] prove the existence of this equilibrium in the Sir Philip Sidney game, a model of signalling among relatives that involves relatedness between signaller and receiver, requiring additional parameters and a more byzantine form than the game considered here.
[2]The game is equivalent in the sense of having the same equilibria.
[3]See Huttegger & Zollman [5] for technical details in the case of the Sir Philip Sidney game; they easily carry over to this case. Further technical details are provided in Hofbauer & Sigmund [25].

References

1. Zahavi A. 1975 Mate selection: a selection for a handicap. *J. Theor. Biol.* **53**, 205–214. (doi:10. 1016/0022-5193(75)90111-3)

2. Grafen A. 1990 Biological signals as handicaps. *J. Theor. Biol.* **144**, 517–546. (doi:10.1016/S0022-5193(05)80088-8)

3. Maynard Smith J. 1991 Honest signaling, the Philip Sidney game. *Anim. Behav.* **42**, 1034–1035. (doi:10.1016/S0003-3472(05)80161-7)

4. Bergstrom CT, Lachmann M. 1997 Signalling among relatives. I. Is costly signalling too costly? *Phil. Trans. R. Soc. Lond. B* **352**, 609–617. (doi:10.1098/rstb.1997.0041)

5. Huttegger S, Zollman KJS. 2010 Dynamic stability and basins of attraction in the Sir Philip Sidney game. *Proc. R. Soc. B* **277**, 1915–1922. (doi:10. 1098/rspb.2009.2105)

6. Borgia G. 1993 The cost of display in the non-resource-based mating system of the satin bowerbird. *Am. Nat.* **141**, 729–743. (doi:10.1086/285502)

7. Caro TM, Lombardo L, Goldizen AW, Kelly M. 1995 Tail-flagging and other antipredator signals in white-tailed deer: new data and synthesis. *Behav. Ecol.* **6**, 442–450. (doi:10.1093/beheco/6.4.442)

8. Chappell MA, Zuk M, Kwan TH, Johnsen TS. 1995 Energy cost of an avian vocal display: crowing in red junglefowl. *Anim. Behav.* **49**, 255–257. (doi:10. 1016/0003-3472(95)80177-4)

9. Horn AG, Leonard ML, Weary DM. 1995 Oxygen consumption during crowing by roosters: talk is cheap. *Anim. Behav.* **50**, 1171–1175. (doi:10.1016/0003-3472(95)80033-6)

10. Gaunt AS, Bucher TL, Gaunt SLL, Baptista LF. 1996 Is singing costly? *Auk* **113**, 718–721. (doi:10.2307/4089004)

11. McCarty JP. 1996 The energetic cost of begging in nestling passerines. *Auk* **113**, 178–188. (doi:10.2307/4088944)

12. Haskell DG. 1999 The effect of predation on begging-call evolution in nestling wood warblers. *Anim. Behav.* **57**, 893–901. (doi:10.1006/anbe.1998.1053)

13. Silk JB, Kaldor E, Boyd R. 2000 Cheap talk when interests conflict. *Anim. Behav.* **59**, 423–432. (doi:10.1006/anbe.1999.1312)

14. Grose J. 2011 Modelling and the fall and rise of the handicap principle. *Biol. Philos.* **26**, 677–696. (doi:10.1007/s10539-011-9275-1)

15. Hurd PL. 1995 Communication in discrete action–response games. *J. Theor. Biol.* **174**, 217–222. (doi:10.1006/jtbi.1995.0093)

16. Számadó SZ. 1999 The validity of the handicap principle in discrete action–response games. *J. Theor. Biol.* **198**, 593–602. (doi:10.1006/jtbi.1999.0935)

17. Lachmann M, Számado Ś, Bergstrom CT. 2001 Cost and conflict in animal signals and human language. *Proc. Natl Acad. Sci. USA* **98**, 13 189–13 194. (doi:10.1073/pnas.231216498)

18. Fudenberg D, Tirole J. 1991 *Game theory*. Cambridge, MA: MIT Press.

19. Gibbons R. 1992 *Game theory for applied economists*. Princeton, NJ: Princeton University Press.

20. Wagner E. Submitted. Dynamics of costly signaling.

21. Godfray HCJ. 1991 Signalling of need by offspring to their parents. *Nature* **352**, 328–330. (doi:10.1038/352328a0)

22. Zahavi A, Zahavi A. 1997 *The handicap principle: a missing piece of Darwin's Puzzle*. New York, NY: Oxford University Press.

23. Bergstrom CT, Lachmann M. 2001 Alarm calls as costly signals of anti-predator vigilance: the watchful babbler game. *Anim. Behav.* **61**, 535–543. (doi:10.1006/anbe.2000.1636)

24. Godfray HCJ. Signalling of need by offspring to their parents. *Nature* **352**, 328–330. (doi:10.1038/352328a0)

25. Hofbauer J, Sigmund K. 1998 *Evolutionary games and population dynamics*. Cambridge, UK: Cambridge University Press.

26. Bergstrom CT, Godfrey-Smith P. 1998 Pure versus mixed strategists: the evolution of behavioral heterogeneity in individuals and populations. *Biol. Philos.* **13**, 205–231. (doi:10.1023/A:1006588918909)

27. Selten R. 1980 A note on evolutionarily stable strategies in asymmetric animal contests. *J. Theor. Biol.* **84**, 93–101. (doi:10.1016/S0022-5193(80)81038-1)