Information and Influence in Sender-Receiver Models, with Applications to Animal Behavior

Peter Godfrey-Smith
City University of New York

Appears in U. Stegmann, (ed.), *Animal Communication Theory*: *Information and Influence.* Cambridge University Press, 2013, pp. 377-396.
The version of the paper here includes a correction to the published version, in footnote 7.

1. Introduction

Debates about animal communication often mirror philosophical debates about communication and meaning in human affairs, especially debates about language.[1] In both contexts there is suspicion in one camp about familiar ways of describing signs in terms of "representation" and the "carrying of information." In some discussions people say that to treat animal communication in this way is "anthropomorphic." But plenty of philosophers would say that these concepts are the wrong ones to apply in the *human* case, as they are part of a discredited picture of human cognitive activity, one that sees us as passively mirroring the world. There is a diverse anti-representationalist tradition in philosophy of language and philosophy of mind.[2]

All options are possible here: an information-based approach might be good in the human case and bad in the animal case. It might be bad in both, or good in both. In an article in the *New York Times* in 2011, Simon Blackburn admiringly notes that biologists studying animal signals have moved beyond a simple representationalist view, and he hopes that philosophers of language will follow their lead. An information-based view might even be bad in the human case and good in the animal case, because the

---

[1] Thanks to Jack Bradbury, Carl Bergstrom, Rosa Cao, and Nick Shea for helpful comments and correspondence. I benefitted also from many other papers in this volume, which I was able to use as a result of Ulrich Stegmann's patience and my colleagues' greater punctuality, for both of which I am grateful.

[2] See Rorty (1982) for a gathering of those threads.

complexities of human language use have overwhelmed a simpler information-carrying role that still exists in animal signaling.

In the animal case, a shift in thinking about communication was linked to a shift in thinking about cooperation. Early work on animal communication was done within a framework that took cooperation and group-level adaptation as common.[3] Some influential criticism of information-based views of signaling, developing from the 1970s, has been associated with a less cooperative view of animal behavior. Dawkins and Krebs (1978) argued that animal communication is a process in which signalers *manipulate* receivers, rather than informing them. A signaler uses another animal as a tool for producing behaviors that help the signaler, regardless of whether the receiver is better off. Given this, it is probably a good idea to "abandon the concept of information altogether" when describing animal signaling (p. 309). Their own later work moderated this view (1984), admitting some role for information and recognizing the role of the receiver as a "mind-reader." Some subsequent work opposed to an informational approach has emphasized that a view based on "manipulation" or "influence" need not be a view according to which receivers are doing badly (Ryan, this volume).

Is a concept of information useful in understanding animal communication at all? If so, is it useful only in understanding cases where there is cooperation? Does the divide between approaches based on information and those based on influence merely reflect a difference in emphasis, without disagreement about underlying processes, or is there a substantive disagreement? If the latter is true, how might further modeling and data collection decide the issue? This paper tries to make progress on all these questions, using a model of sender-receiver relationships that draws on several fields. I argue that within this model, information and influence are complimentary – they come as a "package deal." With respect to the role of cooperation, explaining signaling requires a finer-grained framework than a standard dichotomy between cooperation and conflict, or even a "scale" between them. There are many varieties of *partial common interest*, some of which can stabilize signaling and some of which cannot. In a nutshell, the stabilization of sender-receiver systems ties information and influence together.

---

[3] See Searcy and Nowicki (2006) for a discussion of this historical sequence.

2. Senders and receivers

Suppose there is a *sender* doing something that can be perceived or tracked by another agent, a *receiver,* who acts in a way influenced by what the sender does. Talk of "sending" and "receiving" here is understood minimally, and does not imply anything about information, meaning, or cooperation.

Why are they behaving this way? The question can be asked on many time-scales. The two agents might be making rational moment-to-moment decisions, or inflexibly following behavioral programs shaped by a history of natural selection. But suppose we know that *something* is maintaining this pair of behaviors. Then here is one possibility: the state of the world varies, and acts by the receiver have consequences for both sides. The sender can track the state of the world in a way the receiver cannot. Further, the sender and receiver have *common interest*; they have the same preference orderings over acts that the receiver might produce in each state. For each state they agree about which act is worst, which are better, through to the best. Then a combination of behaviors in which senders produce signals and receivers shape their behavior with the signals can be a *Nash equilibrium*: no unilateral deviation from those behaviors would make either of them better off.[4] If there are enough signals available, perfect coordination can be achieved: in each state of the world a distinctive signal is sent by the sender and it enables the receiver to produce an act tuned to that state.

If a situation like this is found, and provided there is no act that is best in every state, the signals must *carry information*. This formulation has a metaphorical element. Is it really true that information is the sort of thing that can be carried, transferred, or contained? Several critics of information-based views of communication, both in humans and animals, have questioned this point (see the papers by Scott-Phillips & Kirby, Rendall & Owren, and Morton, Coss, & Owings in this volume). The metaphor of containment is indeed probably unhelpful, but what I have in mind can be described in other ways: the state of the signal affects the probabilities of various states of the world.

---

[4] A combination of behaviors by two or more agents is in *weak* Nash equilibrium if no unilateral change by any of the agents would make them better off; the combination is a *strict* Nash equilibrium if every unilateral change would make the agent who made the change worse off. Whether the equilibrium here is strict depends on whether some relevant payoffs are tied.

More exactly, the probabilities of states of the world conditional on the state of the signal differ from the states' unconditional probabilities. This is information in roughly Shannon's (1948) sense. Though this is a "thin" sense of information, many questions arise about it. The use of a physical concept of probability is rightly controversial, and I take its availability for granted. There are also various ways of specifying and measuring the information in a signal. At this stage all that matters is an idea common to many views: in order for the receiver to coordinate acts with conditions in the world, the signal's state must be associated with the state of the world, not independent of it.[5]

The model above is essentially David Lewis' 1969 model of conventional signaling. There is no explanation of how the equilibrium is reached, and Lewis assumed rational choice as the means by which behaviors are maintained. Brian Skyrms (1996, 2010) gave an evolutionary recasting of Lewis' model. "Preference" is replaced by fitness, and rational choice is replaced by a selection process in which present success affects the future frequency of a strategy in a population. Related models have been given in economics.[6] When there is common interest in the sense above, informative signaling can be maintained in a wide range of cases and with various selection mechanisms.

In this kind of signaling, information and influence are coupled together. The signal must be one that carries information about the state of the world, or there is no reason for the receiver to attend to it. The signal must also be one that influences the receiver's actions, or there is no reason for the sender to send it. *Whether the signals carry information about the world* is up to the sender. It is the sender's choice, the outcome of evolution affecting the sender's side, or perhaps a matter of constraint with respect to the properties of the sender. For whatever reason, the sender is producing signals that vary, and that do so non-randomly with respect to some condition in the world. *Whether the signals have influence* is up to the receiver. It is the receiver's choice, the outcome of evolution affecting the receiver's side, or a matter of constraint with

---

[5] For discussion, see Dretske (1981), Skyrms (2010), Godfrey-Smith (2012), and Millikan's contribution to this volume.

[6] For a survey of relevant ideas in economics, see Farrell and Rabin (1996). Throughout this discussion unless otherwise noted, no special costs for signaling are assumed; this is a "cheap talk" model in Farrell and Rabin's sense.

respect to the properties of the receiver. The receiver is acting in a variable way, and non-randomly with respect to the state of the signal.

Perhaps the stabilization of the behaviors in some case is *not* due to common interest. One side may be constrained in a way that enables the other to exploit it. It may be that the sender is doing something that is so well tuned to the receiver's physiology that it can't be ignored. This kind of constraint is the basis for a view of animal signaling defended by Owren, Rendall, and Ryan (2010). Even if a signal can't be ignored, can't its effects be filtered by the receiver? Owren, Rendall, and Ryan argue that senders tend to have the upper hand in a process of this kind. On the other side, the receiver might be tracking something that the sender cannot help but produce – a scent, or reflected light. Camouflage is an attempt not to signal, in this minimal sense of signaling. Another kind of constraint affecting receivers is illustrated by fireflies of one species that produce mating signals of another species to lure males in to be eaten (Lloyd 1975, Stegmann 2009). Given the importance of mating, a tendency to respond to potentially lethal invitations remains in place. These are all cases where one side, sender or receiver, is subject to a constraint that prevents it from responding optimally to the policies pursued by the other.

Setting aside constraints of this kind, let's contrast the wholly cooperative situation with another. Suppose there is complete *conflict* of interest. In each state of the world, sender and receiver have reversed preferences with respect to actions that might be produced by the receiver. Then any information in the signals will be used by the receiver to coordinate acts with the world in a way that suits them, and that is opposed to the sender's preferences. Similarly, any tendency in the receiver to be influenced by signals of the relevant kind will be exploited by the sender to induce acts that are opposed to the receiver's own interests. Any sensitivity on either side can be exploited by the other. So any equilibrium we see will be one where the sender's "signals" say nothing useful and the receiver is not sensitive to them. At equilibrium, there is no information in any signs and no influence either.[7]

---

[7] **This footnote contains a correction to the published version of the paper (July 2013).** Recent work by Manolo Martinez and myself has found counterexamples to this claim – games where agents have reversed preference orderings over actions in all states but where informative

This second case will be revisited below, but if the argument is accepted for now, some consequences follow. Information and influence go together in these cases. At equilibrium there is both or neither. When there is common interest, if information is present in a signal it will come to have influence by receiver adaptation, and if there is influence then the signals will acquire information by sender adaptation.[8] When there is conflict of interest, any influence will be degraded, and any information will be degraded, too.

This reasoning assumes a kind of symmetry with respect to the control that the processes of sender and receiver adaptation have over the situation. This is a strong assumption that may not be realistic in biological cases. When the assumption does not apply, the result can be ongoing exploitation of one side by another, of the kind discussed in the "constraints" paragraph above. In my discussion here, though, I will assume that adaptation is unconstrained on both sides. My focus will be another part of the Lewis model that surely looks biologically contentious, the assumptions about common interest.

So far I have discussed extremes – what I referred to as "common interest" is really *complete* common interest, and what I called conflict is *complete* conflict of interest. In complete common interest, sender and receiver have the same preference ordering over actions in every state. In complete conflict of interest, sender and receiver have reversed preference ordering over actions in every state. Clearly there are many other cases – many kinds of *partial common interest*. Sender and receiver might agree on what is the best action in every state, but disagree otherwise. They might agree on what is worst, but disagree otherwise. If there are enough acts that are relevant, they might agree on best and worst, but flip some in the middle. They might agree entirely in some states but not in others. Complete common interest and complete conflict of interest are extremes, and there are many "paths" between them, as illustrated in Figure 1.

---

signaling can be maintained at equilibrium. These are distinct from the putative exceptions discussed in section 3 below. See Godfrey-Smith and Martinez, "Communication and Common Interest" (forthcoming) for details. These new cases feature mixed strategies by both agents, a possibility not treated in the models below, but the verbal argument given in this paragraph of the published paper is erroneous.

[8] Here I assume that both sides do want different acts to be performed across at least some states of the world. If the receiver has an act which is best in all states, it does not matter what the sender does and whether their interests are alligned.

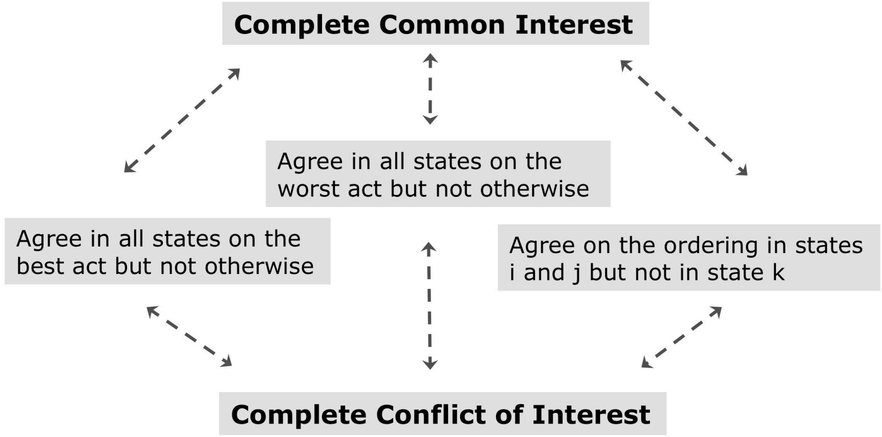

Figure 1: Relations between sender and receiver payoffs. The special cases of complete common interest and complete conflict of interest are linked by many paths that traverse different kinds of partial common interest.

This suggests a hypothesis: informative signaling (when all signals are equally costly to produce) is viable to the extent that there is common interest. In economics, a famous model due to Crawford and Sobel (1982) has this message. That model used a framework where common interest could be measured on a scale. This is not always true; agreeing on the worst-case outcome in every state does not show more, or less, common interest than agreeing on the ordering in most states but disagreeing in others. So it is really only a sketch of a hypothesis to say that informative signaling is viable "to the extent" that there is common interest, but it may be a good sketch. The sketch can be filled out by looking for different ways in which partially informative signaling can be maintained through partial common interest.

3. Model cases with different relations between sender and receiver interests

This section will explore different kinds of common interest, using simple abstract cases. The cases all assume three equally probable states (S1, S2, S3), three acts (A1, A2, A3), and three cost-less signals (m1, m2, m3). The sender perceives the state of world (perhaps its own condition), and can send any of three signals. The receiver perceives the signal and produces any of three actions. None of these steps is affected by "noise" or mis-perception.

The 3-state framework is a good way to look at different kinds of partial common interest, but systems with three states are complicated to analyze; even with *complete* common interest these systems show complex phenomena (Huttegger et al. 2010). My analyses will be very simple. I will describe some equilibria – combinations of sender and receiver policies where each is the best response to the other – and give some dynamic arguments assuming adaptive responses made by one side to a previous move by the other. This is done without considering all equilibria or all possible invaders. In particular, no "mixed strategies" are considered. The only rules considered map states to single messages, and messages to single acts. Where possible, the discussion will be neutral about how the sender and receiver roles are realized – for example, whether senders and receivers evolve in separate but interacting populations, or a single population contains individuals who take both roles in turn.

		States		
		S1	**S2**	**S3**
Acts	**A1**	3,3	2,2	0,0
	A2	2,2	3,3	2,2
	A3	0,0	0,0	3,3

Case 1: Complete common interest

		States		
		S1	**S2**	**S3**
Acts	**A1**	3,3	2,0	0,2
	A2	2,0	3,3	2,0
	A3	0,2	0,2	3,3

Case 2: Agreement on the best act in all states

Tables 1 and 2: The entries in each cell specify sender payoff and receiver payoff, respectively, for each combination of receiver's act and state of the world.

Case 1 is a case of complete common interest. If three messages are available (m1, m2, and m3), there are several ways of using the signals to achieve maximum payoff for both sides. For example, suppose senders send m1 in S1, m2 in S2, and m3 in S3, a policy which can be written as "(S1 → m1, S2 → m2, S3 → m3)." Suppose receivers do A1 in response to m1, A2 in response to m2, and A3 in response to m3, which is (m1 → A1, m2 → A2, m3 → A3). Lewis (1969) and Skyrms (2010) call this outcome a "signaling system" – a combination of policies in which maximum payoffs are achieved by both sides through the use of signals to achieve the best possible match between acts and states. Senders and receivers are then in a strict Nash equilibrium; if either side deviates unilaterally their payoff is reduced. There are six different ways of using the three available messages to achieve a signaling system.

When there is complete common interest, signaling systems can be found by various evolutionary processes. Evolutionary dynamics are complicated in a 3-state case, with other equilibria present and the behavior of the system depending on detailed assumptions about the process. All that matters here, though, is that complete common interest does allow signaling systems to evolve and remain stable.

What effect does moving to *partial* common interest have? First we can note a shift away from complete common interest that does not make a difference to the stability of a signaling system. Suppose the two parties agree on what is the *best* action in all states, but not otherwise. This is case 2. Then the same sender's and receiver's rules that achieved a signaling system in case 1 can again be used to achieve a strict Nash equilibrium. Any deviation from such a system harms the deviator, though how *much* harm any given deviation brings will differ for sender and receiver. (For example, a move to A2 in S1 harms the receiver more than it harms the sender, but it does harm them both.) The changes to the payoff matrix from case 1 to 2 may affect how likely a system is to find this equilibrium, but once a signaling system has been found it is stable.

Next, suppose that sender and receiver agree about the *worst* act in each state, but do not agree otherwise. There are various ways this could happen, and δone is shown as case 3. Another possibility is that the parties could agree entirely about what is good in two states, but disagree in the third – this is seen in case 4.

		States		
		S1	**S2**	**S3**
Acts	**A1**	0,0	2,3	2,3
	A2	2,3	0,0	3,2
	A3	3,2	3,2	0,0

Case 3: Agreement on the worst act in all states

		States		
		S1	**S2**	**S3**
Acts	**A1**	3,3	3,0	0,0
	A2	2,2	2,2	2,2
	A3	0,0	0,3	3,3

Case 4: Complete agreement in two states

Tables 3 and 4: Cases of partial common interest that support informative signaling

I will use the phrase *fully informative signaling* for cases where the sender maps states to signals in a one-to-one manner. In cases 3 and 4, fully informative signaling is not an equilibrium. However, in both cases there is an equilibrium in which *some* informative signaling goes on. In the terminology of Bergstrom and Lachmann (1998), there is a "partial pooling" equilibrium, in which the sender uses the same signal in two of the states, but sends a different signal in the third. In case 3, the following combination of policies is a weak Nash equiibrium: sender does (S1 → m1, S2 → m2, S3 → m1); receiver does (m1 → A2, m2 → A1, m3 → A1). Case 4 has a similar equilibrium: sender does (S1 → m1, S2 → m1, S3 → m2); receiver does (m1 → A2, m2 → A3, m3 → A3).[9]

These are situations with different kinds of partial common interest (PCI), and both have partially informative signaling. As noted earlier, there are different ways of measuring the informational content of a signal, but several of them agree with something like this: if a signal reduces the possible states of the world from three to two, it carries less information than a signal that reduces the possible states from three to one. In case 3, m1 excludes the possibility of S2. It can be seen as having the informational content *S1 or S3 obtains*, whereas m2 has the informational content *S2 obtains*. Because m2 rules out more states, it has more informational content than m1.[10] So at equilibrium, there are *fewer* signals used in these cases and one of the signals carries *less* information than is seen in the equilibria described in cases 1 and 2.

Some general arguments can be given for cases where there are n states, n messages, and n acts. Suppose fully informative signals are sent by the sender, there is

---

[9] These examples are influenced by a case due to Skyrms (2010, p. 81). In Skyrms' case the two parties agree on the worst act in every state *and* agree entirely in one of the three states. The phenomena are separated here. The method used to find the equilibria in the text was to assume a starting point in which the sender uses a one-to-one mapping of states to messages, note the receiver's "best response" to this policy, and continue allowing each side to adapt in turn until an equilibrium was reached. Note that in both equilibria given here, the receiver has a policy for m3 even though this signal is never sent by the sender. These do matter because some receiver's rules for unseen signs open up adaptive possibilities for the sender.

[10] See Godfrey-Smith (2012) for more detail on some of these measures.

disagreement on the best act for at least some states, neither sender nor receiver has a "tie" for their top-scoring acts in any state, and the receiver does not have an act which is top-scoring for more than one state.[11] Then no matter what the receiver is doing, either sender payoffs or receiver payoffs (or both) are not maximal. If receiver payoff is not maximal, there is some change that takes the receiver to their maximum payoff. If, on the other hand, sender payoff is not maximal, then as long as the receiver maps some signal to each act, there is some way of changing the mapping of states to signals that induces the receiver to act in a way that delivers maximum payoff to the sender.

Things are not as simple if the receiver has an act that is optimal in more than one state. Then if the sender sends fully informative signals, the receiver can in some cases achieve maximum payoff by mapping two signals to a single act, and the sender cannot improve their outcome by changing their sending policy, though they can abandon fully informative signaling without penalty. Outside this special case, and the case of *agreement on the best*, partial common interest cannot, in the model used here, maintain fully informative signaling.

I said earlier that informative signaling cannot be maintained when there is complete conflict of interest. This claim has been expressed frequently, but there are some phenomena that can seem to be at odds with it. I will discuss two.

Bradbury and Vehrencamp (2000) present a model that embeds signaling within a general treatment of information use. A receiver is assumed to have various sources of information about a situation of uncertainty, and these establish a "default" behavior. The question the model asks is when attending to a given signal should override this default. And given the effects that the signal will have on the receiver's action, is it worth a sender *sending* the signal, or is the sender better off letting the receiver stick to their default behavior? Bradbury and Vehrencamp discuss four relationships between sender and receiver interests, including what I call complete common interest, complete conflict, and two kinds of PCI. The surprising result is that it is possible for signaling to be beneficial to both sender and receiver in a situation of complete conflict of interest: "senders that

---

[11] The mention of "ties" is included because there is a sense in which sender and receiver can "disagree about the best" without there being incentive to change, if one party gains equal and maximum payoffs from two acts in a given state, while the other gets maximum payoff from only one of those acts.

completely disagree with receivers about appropriate receiver decisions may still benefit by providing moderately honest and accurate signals" (p. 259).

The result can be presented with their hypothetical example, which involves territorial behavior in a mantis shrimp. The owner of a burrow is usually not moulting but occasionally is. When not moulting, a resident will win a contest with an intruder. When moulting, the resident will lose. Intruders have to decide whether to attack or pass by. The question for a resident is whether to threaten intruders, in a way that is a function of whether the resident is moulting. Suppose a resident does threaten in a way that conveys some information about its moult status; it is about twice as likely to threaten when not moulting as when moulting. Once it does this, in effect there are two signals, "threaten" and "meek." It is easy to see that this signaling may be beneficial to the receiver, an intruder. The information about moult status may be sufficiently reliable for the intruder to adopt a policy of passing by if threatened, attacking if not. But this signaling may also be worthwhile to the sender, the resident, because it is not so reliable that intruders do not make mistakes – mistakes they would not make if there was no signaling. If an intruder attacks when the resident is not moulting, the intruder is badly injured and local competition is reduced for the resident. This is complete conflict of interest: in a non-moulting state, the resident is better off if the intruder attacks and intruder is better if he does not. In a moulting state, the resident is better off if intruder does not attack and the intruder is better if he does. The model uses some payoff assumptions that might be questioned, but suppose we grant the assumptions of the case.[12] It seems then to be a situation where informative signaling is viable despite complete conflict of interest.

With respect to the criteria used in this paper, it is not such a case, however. The comparison that Bradbury and Vehrencamp address is this: is signaling of a given reliability better than not signaling at all, for both sender and receiver? This leaves open the question of whether signaling that passes this test will be stable against relevant invasions and modifications, a question their model is not intended to address. Suppose signaling of the kind described in the model is operating, so intruders attack if and only if they see the "meek" behavior from the resident. Then given that a non-moulting resident

---

[12] In Bradbury and Vehrencamp's scenario, the summed payoff to both sides is always higher if there is a fight than if there is not.

would prefer that intruders attack, non-moulting residents should modify their behavior to threaten less, and let receivers attack more often. Given the receiver's present dispositions, the sender benefits from *never* threatening, in fact. Once they do this, signaling has collapsed and receivers will respond by no longer paying attention.[13] The Bradbury and Vehrencamp model describes necessary, but not sufficient, conditions for stable signal use.

There is a second possible challenge to the idea that signaling is not viable with complete conflict of interest. I said that in this situation the only equilibrium can be one where the sender pays no attention to the world and the receiver pays no attention to the sender. But this equilibrium might not be reached. Suppose we have a model like the ones above, with 3,2,0 payoffs possible in each state for each player, complete reversal of preferences in each state, and we start from a situation of fully informative signaling. Assume that each player changes their policy in turn, choosing a "best response" to the present policy of the other player. In at least some cases this produces cycles, where each player exploits and is exploited in turn. Each message is produced in a restricted set of states of the world, but the mapping is always changing.[14] It might be thought that this behavior is an artifact of the simple framework used here, but Elliott Wagner (2011) has a model using a more sophisticated framework that yields a similar result in a situation of complete conflict of interest. His model generates chaotic change, where signals retain informational content at every stage.

In a sense, these are situations in which "informative signaling is maintained." At each moment in time, there is some restriction on the states of the world that can be predicted to obtain, based on the state of the signal. But the dynamic operating does not maintain a particular mapping between signals and states; the mapping is always changing, as the sender adapts to what the receiver has just become.

---

[13] In the framework of an "ESS" model, a population of shrimp who each use the sender and receiver strategies described by Bradbury and Vehrencamp (switching between roles according to their situation) could be invaded by a mutant which always gives the "meek" display when sender, and follows the prevailing rule for receivers.

[14] The length of the cycle depends on the payoff details. There are also "agree on the worst" cases, similar to my case 3, that can cycle in this way.

Situations like this also motivate a distinction between informational content and a concept of *functional* content (Shea, Godfrey-Smith, and Cao, in preparation). Where informational content has to do merely with the way the state of the signal predicts the state of the world, one way to understand functional content is in terms of which states of the world (if any) figure in the explanation of the stabilization of the sender and receiver policies with respect to that signal. In some cases functional content and informational content coincide, and in others they diverge. Many animal alarm calls prompted by flying objects are very unreliable (Searcy and Nowicki 2006, p. 66), being caused most often by non-dangerous birds and other things. When a high "false alarm" rate is adaptive, the informational content of an alarm call may be *predator or harmless bird or...*, even though its functional content is *predator*. This distinction can also be used to describe cycles and chaos due to complete conflict of interest. In those cases signals have ever-shifting informational content, and no functional content at all.

4. Applications

In this section I will apply these ideas to some empirical cases. I do this cautiously, as applying idealized models to real-world systems is a scientific craft acquired by immersion in the details of such cases. A simple case with particular relevance, though, is signals by prey animals to predators. One example is "stotting" by gazelles – stiff-legged vertical leaps, performed on seeing a predator and while moving away. The current consensus is that at least some of these displays are directed at the predators, rather than other members of the prey species, and a variety of animals use displays of this kind (Searcy and Nowicki 2006, Bradbury and Vehrencamp 2011).

Searcy and Nowicki describe these cases as follows: "If signaling is from prey to predator, then signaling is between individuals with interests that are diametrically opposed, rather than convergent" (2006, p. 53). Clearly predator and prey have opposing interests in many respects, but this is not a case of complete conflict. There are some states of the world in which their preferences for predator behavior match, because some outcomes of an interaction are bad for both sides. A payoff matrix is given in Table 5: sender and receiver agree in one state of the world but not the other. If the prey animal is healthy and strong enough to escape, a chase is undesirable for both sides. If the prey is weak, a chase is good for predator and bad for prey.

		States	
		Strong	**Weak**
Acts	**Chase**	2, 2	0, 5
	No chase	3, 3	3, 3

Table 5: Predator and prey

A case like this is also discussed by Bradbury and Vehrencamp (2000), using the model described earlier. They show that if a signal is sufficiently reliable, both sides can do better with signal use than without it. This is another case with a possibility of adaptive degrading of the signal's informativeness by the sender, however. Once receivers do not chase an individual that sends a certain signal, prey animals weak enough to be caught will benefit from sending this signal. Senders will degrade the information content of signals, and receivers will adapt by paying no attention to them. This suggests that given that such signals do persist, additional factors are needed to explain them (Fitzgibbon and Fanshawe 1998). In the gazelle case, a likely extra factor is that weak animals can rarely manage an impressive stott. Stotting is a signal of vigor that is hard to fake.

Some other signals apparently directed at predators do not seem to have this feature. One is "tail-flagging" by white-tailed deer (*Odocoileus virginianus*), in which the deer raises its tail to reveal white patches on its rump and the underside of the tail itself (Caro 1995). Another is seen in some kinds of squid (eg., *Sepioteuthis*), which display with color changes and elaborate arrangements of arms when predatory fish approach (Moynihan and Rodaniche 1982). These do not seem to be displays of vigor, like stotting, and may just be signals to the predator that they have been perceived.[15]

Evidence that these other signals actually deter predators is often weaker, but let's assume that they do. Why would the informativeness of these signals not degrade? Once an "I've seen you" display deters predation, why not maintain it all the time, whether you

---

[15] Hanlon and Messenger (1996) note that in the case of squid, the display may have the function of startling the predator and disrupting the behavioral routines it employs in predation.

have seen a predator or not? Bradbury and Vehrencamp (2011, Chapter 14) note that some predator displays emphasize that the prey is attending in exactly the right direction, which makes dishonesty difficult. Another factor may also become relevant. In the squid case, it is not that these displays would be difficult to fake, but they probably use energy and interfere with other behaviors. If predators are in fact rare, there may be no sense in disrupting other activities to perform the display when a predator has not been seen. The economics of the display, from this point of view, are no different from the economics of hiding, or running away. So at least to some degree, an "I have seen you" signal will naturally be reliable, as it will only be when a predator has been seen that the signal is worth producing. The signal is not *hard* to produce falsely but *pointless* to produce falsely. This argument relies on the assumption that the display does in fact disrupt other activities or use resources. In the case of squid displays this seems likely, but in the case of tail-flagging by deer it perhaps seems less so. Bergstrom and Lachmann (2001) give a model in which yet another reason for the stability of predator-directed signals is suggested, that may also apply to the squid case and perhaps to the deer. When an individual produces a signal directed at predators when none has been perceived, it risks alerting unseen predators to the individual's presence and location. These further explanations for the maintenance of honesty in the signal are complements, not alternatives, to the explanation in terms of partial common interest; it is partial common interest that makes the signals worth producing at all, and other factors that prevent the system being undermined by dishonest senders.

Another situation with partial common interest is mating signals. A standard picture of these cases is as follows (Searcy and Nowicki 2006, Ch. 3). In many though not all cases, males are the main signalers and females are receivers. This is because a female would like to mate if a male is of high quality but not otherwise, while males are less choosy. So males will advertise, and the females' problem is sorting the advertisements and choosing good males. This is a case of partial common interest, as both parties would prefer to mate if the male is high quality, and the problem of maintaining the informativeness of signals is also clear. If females respond to indicators of quality that can easily be faked, then they will be faked, and females will stop attending to them.

While accepting the assumptions of the standard picture for purposes of discussion, further factors can be noted. The range of relevant states of the world does not only include high and low quality states of the male. Male signals often achieve other reductions of uncertainty, by indicating that the sender is of appropriate species, age, and sex, along in some cases with location. Neither side has an interest in attempting a mating with someone of the wrong species, age, or sex: to that extent there is common interest. Against that background, there are states of the world where common interest breaks down, when a female discriminates between high and low quality males. This divergence motivates receivers to attend to displays of quality that are hard to fake, or too costly for low-quality senders to produce. So the overall story has a "layered" structure. Common interest explains why signaling is done at all, and the particular form that signaling takes is due to selection by the receiver of signals whose information content is hard to degrade. If receivers in some cases did *not* have a way of enforcing use of a hard-to-fake indicator of quality, and had to deal with use of a simple cheap-talk signal of availability by senders, then we would have a PCI payoff matrix of the kind discussed earlier, where the two parties share preferences in two states of the world and diverge in one. Depending on the assumptions made about payoff relations and probabilities of states, receivers may accept the refusal of senders to distinguish low and high quality states, and use the less informative signals that result to guide their mating behavior, or they might ignore the signals and use other cues.[16]

In sum: in sender-receiver systems, common interest maintains informative signaling by connecting information with influence. Without common interest, both information *and* influence should degrade, unless constraints of some kind prevent this, and this they may do; my discussion assumes unconstrained and comparable processes of adaptation on the sender and receiver sides. A contrast might be drawn between the

---

[16] For example, suppose the payoffs for sender and receiver respectively if a mating occurs are (5,5) in the case of a high quality male, (5,0) if low quality male, and (1,1) if the male is of the wrong species, while the payoffs if no mating occurs are (2,2) regardless of the state of the male. Then if the system starts with unique signals being sent in each state and the receiver making their best response, senders will "pool" two states into one signal, obscuring the distinction between low and high quality males. If states are equiprobable then receivers should continue to respond to the signals.

human and animal cases. In animal signaling, where change is slow and often involves genetic evolution, much of what might compromise the application of these simple adaptationist models is constraint on adaptation. In the human case, the problem comes from our multifarious interests and capacities – from our flexibility. Returning to the model: common interest is not a yes-or-no matter, but is structured with two extreme cases (complete common interest and complete conflict) linked by many intermediates. One aim of this paper is to supplement talk of "cooperation and competition" with this finer-grained framework. Complete common interest may be rare in a between-organism context, but partial common interest is not, and various kinds of partial common interest support partially informative signaling, even in situations where talk is free.

References

Bergstrom, C. & Lachmann, M. (1998). "Signaling among relatives. III. Talk is cheap." *Proceedings of the National Academy of Sciences USA* 95: 5100–5105.

Bergstrom, C. & Lachmann, M. (2001). "Alarm calls as costly signals of antipredator vigilance: the watchful babbler game." *Animal Behaviour* 61: 535-543.

Blackburn, S. (2011). "Of Hume and Bondage." *New York Times*, December 11, 2011.

Bradbury, J. & Vehrencamp, S. (2000). "Economic models of animal communication." *Animal Behaviour* 59: 259–268.

Bradbury, J. and Vehrencamp, S. (2011). *Principles of Animal Communication, 2nd edition*. Sunderland: Sinauer.

Caro, T. (1995). "Pursuit deterence revisited." *Trends in Ecology and Evolution* 10: 500-503.

Crawford, V. P., & Sobel, J. (1982). “Strategic information transmission.” *Econometrica* 50: 1431-1451.

Dawkins. R. & Krebs. J. (1978). "Animal signals: information or manipulation?" In J. Krebs and R. Davies, eds., *Behavioural Ecology: an evolutionary approach*. Oxford: Blackwell, pp. 282–309.

Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press.

Farrell, J. & M. Rabin (1996). "Cheap talk." *Journal of Economic Perspectives* 10: 103-118

Fitzgibbon, C. D. & Fanshawe, J. H. (1988). "Stotting in Thomson’s gazelles: an honest signal of condition." *Behavioral Ecology and Sociobiology* 23: 69–74.

Godfrey-Smith, P. (2012). Review of Brian Skyrms' *Signals* (2010), forthcoming in *Mind*.

Hanlon, R. & Messenger, J. (1996). *Cephalopod Behaviour*. Cambridge: Cambridge University Press.

Huttegger, S., Skyrms, B., Smead, R., & Zollman, K. (2010). "Evolutionary dynamics of Lewis signaling games: signaling systems vs. partial pooling." *Synthese* 172: 177–191.

Krebs, J. & Dawkins, R. (1984). "Animal signals: mind-reading and manipulation." In J. Krebs and R. Davies (eds.), *Behavioural Ecology: an Evolutionary Approach*, 2nd ed. Oxford: Blackwell, pp. 380–402.

Lewis, D. K. (1969). *Convention*. Cambridge, MA: Harvard University Press.

Lloyd, J. E. (1975). "Aggressive mimicry in *Photuris* fireflies: signal repertoires by femmes fatales." *Science* 187: 452–453.

Milikan, R. G. (this volume). "Natural information, intentional signs and animal communication."

Morton, E., Coss, R., & D. Owings (this volume). "Mitogenetic rays and the information metaphor."

Moynihan M. & Rodaniche A. (1982). "The behaviour and natural history of the Caribbean reef squid *Sepioteuthis sepioidea* with a consideration of social, signal and defensive patterns for difficult and dangerous environments." *Advances in Ethology* 125: 1–150.

Owren, M., Rendall, D., & Ryan, M. (2010) "Redefining animal signaling: influence versus information in communication." *Biology and Philosophy* 25: 755–780.

Rendall, D. & Owren, M. (this volume). "Communication without meaning or information: Abandoning language-based and informational constructs in animal communication theory."

Rorty, R. (1982). *Consequences of Pragmatism*: *Essays 1972-1980*. Minneapolis: Univesity of Minnesota Press.

Scott-Phillips, T. & Kirby, S. (this volume). "Information, influence and inference in language evolution."

Searcy, W. and Nowicki, S. (2006). *The Evolution of Animal Communication.* Princeton: Princeton University Press.

Shannon, C. (1948). "A mathematical theory of communication." *The Bell System Mathematical Journal* 27: 379-423.

Shea, N. Godfrey-Smith, P. & Cao, R. (in preparation). "Content in simple signaling systems."

Skyrms, B. (1996). *Evolution of the Social Contract*. Cambridge, MA: Cambridge University Press.

———. (2010). *Signals: Evolution, Learning, & Information.* New York, NY: Oxford University Press.

Stegmann U. (2009). "A consumer-based teleosemantics for animal signals." *Philosophy of Science* B: 864–875.

Wagner E. (2011)."Deterministic chaos and the evolution of meaning." *The British Journal for the Philosophy of Science*. Online ahead of print at doi: 10.1093/bjps/axr039