February 2010

Signals, Icons, and Beliefs

Peter Godfrey-Smith

Harvard University

Will appear in D. Ryder and J. Kingsbury, and K. Williford (eds.),

Millikan and Her Critics. Blackwell, pp. 41-58.

1. Introduction

One way to approach the explanation of meaning and semantic content is to hold that any object which is a sign or a representation – anything that has semantic properties – has this status because of the way it is used in an interaction between two other kinds of things, a *sender* and a *receiver*. "Sending" is understood here in a very broad way, to include any kind of creation, display, or inscription of a sign. "Receiving" involves the sign's interpretation and use.

A defender of this approach has to say something about how semantic phenomena that do not apparently involve senders and receivers can be understood. If our thoughts have content, for example, that does not seem to be because they are being communicated from one place to another. The defender of the sender-receiver view might respond by retreating from the attempt to cover all cases, or they might claim that their framework is more general than it looks.

There are also very different approaches to explaining content. We might start with an explanation of the content of thoughts and explain other semantic phenomena in terms of the *expression* of these thoughts. Such an approach might be motivated by the view that anything that deserves to be called a "sender," "receiver," or other user of a message must itself have states with semantic or intentional properties. Perhaps only thoughts have "intrinsic intentionality," although various other signs (maps, spoken sentences, etc.) have "derived intentionality" as a result of their use by thinkers (Searle 1992).

Ruth Millikan's theory is the most carefully worked-out version of the sender-receiver approach to semantic phenomena within naturalistic philosophy – perhaps within any kind of philosophy. The element of Millikan's view that has caught the most attention is the "teleological" element, the use of a biological notion of function, grounded in history, to explain meaning. Less discussed is the fact that hers is an ambitious version of the sender-receiver approach; any entity that has semantic content, according to Millikan's analysis, has it as a consequence of its relations to a "producer" on one side and an "interpreter" or "consumer" on the other. So if our thoughts have content, there must be "producer and consumer" arrangements inside us, as well as between us. The same schema is used when explaining the semantic properties of external signs and internal states, and neither has semantic properties that are prior to or reducible to the other.

An additional perspective on sender-receiver systems has recently been provided by Brian Skyrms, in *Signals* (2010). Skyrms takes a modeling approach, looking at how sender-receiver configurations come into existence and are maintained, especially in situations where foresight and deliberation are not involved. Recent years have also seen much new work on this topic within evolutionary biology.[1] Bringing these ideas together opens up new avenues for the development of the sender-receiver approach. This paper pursues a synthetic project of that kind. The next section describes sender-receiver configurations. I then compare some explanations of semantic content within that framework. The final section looks at the application of the sender-receiver approach to mental states and internal representations.

---

[1] See Maynard Smith and Harper (2003), Lachman et al. (2001), Rendall et al. (2009).

2. Senders and Receivers

Vervet monkeys give alarm calls which alert their troop when a predator is spotted (Seyfarth, Cheney, and Marler 1980). The hearers use the calls as a guide to the state of their environment, and the vervets give different alarms when they see different predators (leopards, snakes, eagles). Paul Revere used an alarm signal during the American Revolution ("one if by land, two if by sea"). Actions directed on the British army were guided by attending to a lantern in a church. Obvious features present in these cases include a sender, a signal, and one or more receivers. The signal is produced by the sender, and this is done in a way that is responsive to the state of something in the world beyond the signal itself. A receiver produces behavior according to what the signal is like. These behaviors are also directed on something beyond the signal – often, on the state of the world that the sender was responding to when producing the signal.

Sender-receiver configurations have been treated as a starting point in discussions of meaning within very different traditions.[2] This is also the set-up that provides the starting-point for the mathematical theory of information and communication, and was pictured by Claude Shannon (1948) in a famous diagram (Figure 1). Shannon's diagram adds some elements beyond the basic sender-receiver structure described above, and omits others. He includes a source of "noise," for example, and aside from the signal itself there are extra "messages" on each side. Shannon says that the "destination" on the far right is "the person (or thing) for whom the message is intended."

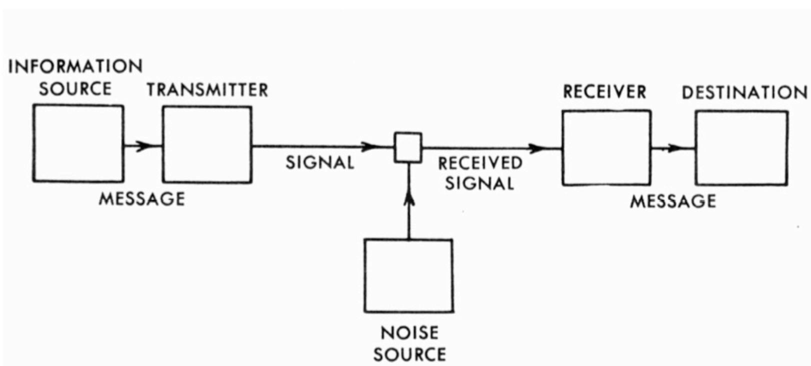

Figure 1. Shannon's diagram of a "general communication system"

---

[2] See, for example, Eco (1976).

Many different terminologies are used. Shannon's picture has a "transmitter" and a "receiver." The game-theoretic literature tends to say "sender" and "receiver." In Millikan's framework we have "producers" on one side and "interpreters" (1984) or "consumers" (1989, 2004) on the other. Between the two we may find "signals," "symbols," "representations," or "signs." Millikan's general term for a sign-like object is "icon." She reserves "representation" and "signal" for narrower categories. In this paper I will move between different terminologies in a way that is not intended to reflect theoretical differences, and the aim is to discuss a model that can be applied very broadly. A sender creates or emits things that can be called "signs" or "signals," and they are read or interpreted and put to use by a receiver. The term "reader" will also be used, as it emphasizes the idea that what a receiver does is apply an interpretation rule to the sign. When discussing Millikan's work I will mostly use her preferred terms. Millkan's term "producer" is in some ways better than "sender," as it emphasizes that a sign is being created as well as transmitted or transported. Her "consumer" is perhaps *too* broad in its connotations, and that will become an issue in the last section of the paper. But I will not impose a standardized terminology, and will use whichever terms have the most useful connotations for the case being discussed.

As the sender-receiver approach sees things, just about anything can function as a sign, and can have any semantic content, if it has an appropriate sender and receiver. A receiver should not be mistaken here for an interpreter in the sense of a mere onlooker, someone trying to "make sense" of the sign. Senders and receivers have to be tied together more closely than that, as will be discussed below. A derivative case may be recognized, where an interpreter makes use of "natural signs" or "cues" to guide action. (Smoke is a natural sign of fire.) These are produced in a way that involves no coordination with the interpreter. These signs can carry information in the sense described by information theory - anything which changes the probability of some other event in the world carries information about it. They might also be described in other quasi-semantic terms, perhaps through a pretence that they have an appropriate sender. But they are not taken to have content in the same way that a sign within a sender-receiver configuration does.

When we fix the sender-receiver configuration in our minds and go looking through nature, we find a great diversity of things that fit the pattern. Some are very simple, some complex. Sometimes there is a single receiver, sometimes several. Many are found between whole organisms: alarm calls, mating calls, chemical signals that mark trails and territories, and signals used by bacteria to work out how many of them are in the same area ("quorum sensing"). We also find sender-receiver configurations within single organisms; hormones (like adrenalin, thyroxin, and testosterone) are examples of such internal signals.

Signaling within certain tightly integrated colonies and societies is a special case. Bee dances, which Millikan has discussed in detail, provide an ideal example. When a honey-bee finds a source of nectar that is fairly far from the hive, she will return to the hive and perform a looping dance, near a vertical surface, which includes a "waggle" in the middle. The dance signals the location of the nectar source. The dance's angle off the vertical corresponds to the angle between the nectar and the point where a line drawn down from the sun hits the horizon, in relation to the hive. The length of the "waggle" segment of the dance represents the distance. Worker bees, which are normally sterile, have tightly connected evolutionary interests. Some biologists see honey bee colonies as "superorganisms," with individual bees akin to cells (Holldobler and Wilson 2008). Then signaling between individual bees is more like signaling between parts of a single body. Maybe the bee dance is best seen as intermediate between a case of signaling between multiple individuals and signaling between the parts of one. And when we look closer at some "organisms," including familiar trees such as an oak, we find that they are in some ways like collections of individuals, not single ones (White 1979). When there is signaling within a plant (for example, using the hormone auxin), this takes place within a collection that is more physically integrated than the bees in a colony, but with less reproductive specialization – all the branches of an oak tree have independent reproductive capacities. So there is a continuum between cases with separate agents who come together and signal in their own interests, and cases involving signaling within one agent. Here it is interesting to note that some neurotransmitters are thought to have evolved from chemicals whose initial role was inter-organism signaling.

There are also unclear and questionable cases of the sender-receiver configuration. DNA looks very much like it was designed to send information (Shea 2007, Bergstrom and Rosvall forthcoming). But who is sender and who is receiver? "Messenger RNA" looks like a message sent *from* the genes to the machinery of protein synthesis, but who is the sender in the case of genes themselves? Even with the questionable cases set aside, we have a large and diverse collection: human language, subway maps, bee dances, alarm calls, birdsong, hormones, pheromones.... The sender-receiver configuration is found at many different scales and has arisen independently many times. So now we can ask: why should we find things in this arrangement? Why is it common? What makes it come about, and what keeps it from falling apart?

David Lewis (1969) gave a game-theoretic model of sender-receiver coordination. In a ground-floor version, assume there is a sender who can see the state of the world and produce signals. The receiver can see only the signal, but can act according to the signal he sees. There are two states of the world and two actions available. The sender can send the same signal all the time or adjust the signal to the state of world. The receiver can act the same way all the time or adjust his behavior to the state of the signal. A situation of *common interest* is assumed; for each state of the world there is a right and wrong action, and both agents benefit if and only if the receiver produces the right action. A *signaling system* is a combination of policies by sender and receiver such that both of them do reliably get the payoff. That means that the sender is adjusting its signals to world's state, and the receiver is interpreting these in the right way. A signaling system in this scenario is a *Nash equilibrium* - a combination of behaviors such that neither agent does better by unilateral change. Each behavior is the "best response" to the other. Signaling can be maintained by common knowledge of this fact.

Skyrms (1995, 2010) takes an evolutionary approach. His model does not assume "smart" senders and receivers, and applies to bees and bacteria as well as to more complex agents. In one of his models, we assume a population of individuals who take both sender and receiver roles on different occasions, occupying each role half the time. As in the Lewis model, only senders can see the state of the world but only receivers can act on it. Each individual has a behavioral profile that combines a sending policy and a receiving policy. With two states of the world, two possible signals, and two possible actions, there are 16 such profiles. The agents interact randomly, receive payoffs, and reproduce asexually according to those payoffs. Both agents receive a payoff whenever the receiver does the right act for the state of the world, and otherwise no one is paid. The states of the world are equally likely. Agents reproduce asexually, in a way reflecting their payoff. Then it can be shown that evolution favors behavioral profiles which give rise to signaling systems. Those behaviors are *evolutionarily stable* – they cannot be invaded by rival strategies once they are common. Other behavioral profiles are not evolutionarily stable. Skyrms found in simulations that his idealized populations did make their way from a variety of initial states to signaling systems. In other models, the sender and receiver roles are separate and fixed, there may be multiple receivers, and so on. Adaptive change may also be due to learning dynamics rather than evolution; there are several different processes by which sender-receiver configurations can arise and be maintained.

In the models sketched so far, "common interest" between sender and receiver is assumed. Similarly, Millikan's discussions of "producer-consumer" systems include the requirement that these two devices "cooperate" with each other, as a consequence of evolutionary design. Millikan does not say a lot about what degree of cooperation is required here, and she views the problem of explaining the maintenance of cooperation as over-emphasized in some discussions of social evolution, at least in the case of humans.[3] But clearly it is worth asking how sensitive models of signaling are to assumptions about common interest.

In Skyrms' 2010 models, a total breakdown of common interest dooms signaling – this is what we have when sender and receiver benefit from completely different acts being performed in a given state of the world. But there are also cases of *partially* aligned interests. In the table below, the payoffs are given for sender (first) and receiver (second) in each combination of a state of the world and an act performed by the receiver. For example, if the receiver produces act 1 in state 1, the sender gets 2 units and the receiver gets 10. This is a case where in states 1 and 2 there is a partial conflict of interest: the

---

[3] "Most aspects of social living involve cooperation in ways that benefit everyone" (2004, p. 21). Even if this is true of present-day interactions that maintain some central practices of public language use, especially in Connecticut, it is not something that can be generally assumed in an evolutionary context. In those contexts it is also relative advantage that usually matters.

sender would rather the receiver performed act 3 in states 1 and 2, while the receiver would rather perform act 1 in state 1 and act 2 in state 2. In state 3, their interests align.

	Act 1	Act 2	Act 3
State 1	2, 10	0, 0	10, 8
State 2	0, 0	2, 10	10, 8
State 3	0, 0	10,10	0, 0

Table 1: Partially aligned interests (from Skyrms 2010)

In a system of this kind, there is an equilibrium if everyone uses the following strategy: *if sender, send signal 1 in states 1 and 2, and send signal 2 in state 3; if receiver, do act 3 in response to signal 1, and do act 2 in response to signal 2.* Effectively, senders refuse to discriminate between states 1 and 2, making it rational for the receiver to choose their second-best option. This *second*-best option for the receiver gives the *best* outcome for the sender. The receiver would benefit from getting perfect information about states 1 and 2, but there is no reason for the sender to send it.[4]

There are also situations where full and honest signaling can be maintained even though it has an "altruistic" character. Alarm calls are good cases to think about. Suppose there are two types, a signaler and a "free-rider" which responds to alarm calls but keeps quiet when they are the one who sees a predator. If these individuals spend time foraging in pairs (not mating pairs), the best thing to be is a free-rider in the company of a signaler: they will alert you to predators, incurring some danger by making the call, and you get the benefit of their calls without making any of your own. The next best thing is to be a signaler with another signaler. The next best is to be a free-rider with a free-rider, getting no warnings but taking no special risks, and the worst is to be a signaler with a

---

[4] See Stegmann (2009) for discussion of a different kind of case involving conflict of interest, featuring exploitation of a sender-receiver system by a malicious (predatory) alternative sender.

free-rider. This is a "prisoner's dilemma."[5] In a population containing both types, signalers are always exploited when in mixed pairs, but signaling can prevail provided that interaction is sufficiently *correlated* (Hamilton 1975), so that most signals are sent to other signalers and free-riders are stuck with their own kind.[6]

Above I noted that sender-receiver configurations are found in nature both within and across organisms, and also in cases where the boundary of an organism is not clear. In general, as we move further from the within-organism case, breakdowns of common interest become more likely. They are not impossible within even organisms like ourselves, however. David Haig argues that within our bodies there is the potential for competition between *paternally derived* and *maternally derived* genes. Roughly speaking, the genes an animal gets from its male parent have less interest in the well-being of the animal's siblings than its maternally derived genes do, as long as there is any uncertainty about the paternity of those siblings (Haig 1996). So our paternally derived genes would have us monopolize resources in a family situation more than our maternally derived genes would. This sort of conflict has been well-documented in pregnancy, where the "resources" in question are supplied directly through the mother's own body, and the conflict is played out in a dynamic between senders and receivers in a hormonal signaling system. The extension of the hypothesized conflict beyond pregnancy is, I understand, still more speculative (Haig 2010).

In other within-organism cases, the idea of common "interest" between sender and receiver becomes tenuous and metaphorical. But the parts of an organism can have a "cooperative" relationship if they have evolved to interact with each other in a way that furthers the interests of the whole organism or some other reproducing entity. For example, the most physically tiny signaling systems are probably those within individual bacteria (Camilli and Bassler 2006). A receptor within a cell may function as a receiver of messages sent from its genome, or a relay-like "sender" which transduces an external

---

[5] It is sometimes also required that the payoff from mutual cooperation be less than half the average of the payoff from defecting on a cooperator and the payoff from being a cooperator who is defected on (Axelrod 1984). I assume here that there is no possibility of conditional behavior in response to the behavior of a partner.

[6] The biological literature has also looked extensively at the idea that the *costliness* of certain signals can maintain their honest use in a population. See Maynard Smith and Harper (2003).

cue. The receptor itself cannot reproduce or help other receptors similar to it to reproduce, but the bacteria have evolved to make receptors which interact with other parts of the cell in a way that furthers bacterial reproduction.

Very simple cases like this are also significant because they probe the boundaries between sender-receiver systems and set-ups that should be understood in different terms. There is a tendency in cell biology to use the language of "programs" and "computation" in a very broad way about intricate arrangements that are the products of evolutionary design, and this leads also to descriptions in terms of meaning and intention. But being a sender-receiver system is a quite different matter from (for example) being a "programmed device." In the paradigm cases of sender-receiver configurations, receivers are physically distinct from the signs they interpret, and receivers can act in ways that have consequences for senders and signs as well as themselves. Sender, sign, and receiver are not merely three things connected in a causal chain. In some discussions, it is also said that an essential feature of sign use is a potential flexibility in the interpretation rule applied by a receiver. This is referred to as "arbitrariness" (Maynard Smith 2000) or "conventionality" (Skyrms 2010). Something like this is certainly a feature of familiar cases of public representations used by humans. It is not so clear in some of the biological cases above, where we find mechanically simple receivers applying "hard-wired" rules. Skyrms says that at least "contingency" in the interpretation rule is seen in those cases. I think this intuition is linked to the basic requirement of the separability of sign and reader: each could be independently modified in principle. It also seems that when a sign's effect on a reader is too physically constrained – when the sign's physical impact leaves "no room for interpretation," as will often be the case with the parts within a cell – then we have something that is not a clear case of the sender-receiver configuration.[7]

---

[7] Chapter 7 of Millikan (1984), "Kinds of Signs," surveys other kinds of simplicity and sophistication in sign systems, emphasizing further distinctions between cases that clearly fit the basic requirements described here.

3. Content

Sender-receiver configurations arise often in nature and at many scales. They are also prominent everyday features of our own social lives. As a result, we have culturally entrenched habits of dealing with them and talking about them – habits of description and ascription of content. Using language (a sign system) we constantly talk about what messages in other sign systems "mean" and "say." So there are two sets of facts here. First, there are facts about the relations between signs, their users, and the world that affect how signaling systems work. Second, there are facts about our habitual ways of talking about signs within social life. This second dimension might be seen as a kind of "folk theory" of signs, though it probably has roles in social coordination that are poorly described with the term "theory."

When investigating sender-receiver systems in nature we can expect to be guided by those habits. This might have both good and bad effects: good, as we are sensitive to the presence of signaling systems, perhaps bad, as we may describe naturally occurring sign systems in ways that are guided too much by everyday cases. For example, we might describe the meanings of simple signs in a way that is affected by our experience with the very sophisticated semantic properties found in language.

In this section I will look at the content of signs in sender-receiver systems. People can agree that sender-receiver configurations are the natural home of meaning, but disagree about how the content of messages in such systems should be described. I will start by looking at an approach making use of information theory. This approach was pioneered by Dretske (1981), but here I discuss a version due to Skyrms (2010). I then compare this approach to Millikan's.

On an information-theoretic view, a signal contains information when it changes the probabilities of states of the world. Skyrms' suggestion is that the *amount* of information in a signal is measured by the amount of change the signal makes, on average, to the probabilities of each of the possible states of the world. The *content* of a signal – what it says – is represented with a list of *all* the differences made to the possible states. "Propositional" content – content expressible with a "that..." clause – is for Skyrms a special case. When some states of the world are not just made less likely by a signal but are ruled out, then the content of the signal can be expressed as a proposition which disjoins the states that remain.

Clearly these are real features of any sender-receiver configuration (bracketing foundational questions about probabilities), and they figure in explanations of how the system works. If a signal did not change the probabilities of relevant states of the world there would be no reason to attend to it.[8] The case of partially aligned interests in sender and receiver discussed in the previous section illustrates the possibility of more fine-grained analysis in these terms. When that system is at the equilibrium described earlier, the mixture of matched and mis-matched interests across sender and receiver is reflected in the amount of information contained in different signals. The signal sent in states 1 and 2, when there is some conflict, contains less information than the signal sent in state 3, when there is none.

Not all of the properties that might be recognized by an information-theoretic analysis have this role, however. Skyrms also says that a signal in a sender-receiver system has a second kind of content. As well as changing the probabilities of states of world, a signal changes the probabilities of acts by the receiver. Skyrms says a signal contains information about acts in the same sense in which it contains information about the states of the world. But this sort of "content" does not have anything like the same role as the first. In a sender-receiver configuration, the role played by a signal is to tell the receiver something about the state of the world, not to tell the receiver how the receiver is likely to act. We as onlookers can learn about actions likely to occur by looking at the signals, but this is surely not something the receiver is being told. This is an illustration of a more general fact: any signal will have effects on the probabilities of a huge range of goings-on, but almost all of these are irrelevant to a description of how the signaling system works and why it exists.

---

8 Here I set aside a very special case, where an uninformative signal might act as useful randomizing device in a situation where mixing of behaviors reduces variance in payoff in a way that is evolutionarily advantageous (Godfrey-Smith 1996, Chapter 7). I should also note that Millikan thinks that the foundational questions about probabilities set aside here do raise problems for information-theoretic views.

Information-theoretic views have often had problems dealing with phenomena involving error and false content.[9] Here Skyrms attempts a very simple treatment. He says that when a signal raises the probability of a state which is not actual, that is "misleading information, or *misinformation*" (p. 80). This only provides a very partial analysis, however. To explain the possibility of error or false content is usually seen as a matter of saying how a representation can have a propositional content which is false: how can a signal say that *p* is the case when it is not? For Skyrms, a signal only has propositional content when it rules out one or more states of the world completely. So for a propositional content to be false, a signal must say that some state has been ruled out when in fact it is actual. That is impossible within Skyrms' model. When a non-actual state's probability is raised and the actual state's probability is lowered by a signal, that does give the signal the potential to mislead an interpreter. But in Skyrms' model a signal cannot *say that* something is the case when it is not.[10]

As Dretske found in 1981, there is a certain amount about semantic phenomena that can be explained very readily in an information-theoretic framework, and a residue which is a persistent problem. One response is to add extra elements to the information-theoretic framework, as Dretske (1988) did. Another is to set information theory aside and start again, as Millikan does. Yet another would be to say that some of the recalcitrant "semantic phenomena" are just features of the way we *talk* about representations, features of our interpretive habits. Or these features might be real but found only in most sophisticated sign systems, and should not be read back into simple sender-receiver configurations.

I now turn to Millikan's view. For Millikan, the content of a sign or "icon" derives from the biological functions of its producers (senders) and consumers (receivers), and from the normal conditions under which these functions are performed. Both the

---

9 See Fodor (1984), Godfrey-Smith (1989 & 1992).

10 Skyrms also says that a "misinformation" has been transmitted when a receiver acts differently from how he would if he had "direct knowledge" of the state of the world. If what really matters here is a counterfactual about rational behavior, this compromises the naturalistic status of the analysis and its applicability to simple systems. The idea might be expressed in a more direct way, by saying that when the payoff is low then misinformation has been transmitted. But why, within a non-historical analysis of content like Skyrms', does the size of a payoff occurring *after* the signal has been received matter to what the signal *said*?

functions and the "normal conditions" depend on the system's history. "History" sounds like a long time ago, but the relevant history is made up of the processes of maintenance and stabilization discussed throughout this paper.

In simple cases like animal danger signals and bee dances, Millikan thinks that an icon has two kinds of content – or, more accurately, a kind of content that has a dual character when we express it in language. A bee dance says both "There is nectar at location X" and "Go to location X." There is both a descriptive element (indicative) and a command (imperative). Only in more complex cases do we find contents that are purely indicative or imperative. Suppose there is no specific outcome which a consumer is supposed to bring about in response to a given icon. The icon is instead supposed to enter into something like a process of inference, with the actions resulting being affected by other icons and other states of the system. In that case the sign will have a purely indicative content. The kind of dual content that Millikan recognizes in simple sender-receiver cases is different from Skyrms', and conforms better with the role that signs have in such systems. The sign does not say which actions have become more likely; it instructs the receiver what to do.

From now on I will mostly discuss the indicative side of the content of these simple signs. This content is not determined by which states of the world are made *likely* by the signal. Instead, the content depends on *how the world has to be* if the icon is to affect the activities of its consumers in a way that leads to them to perform their functions in a historically normal way – in the way that led to that pattern of icon-use being favored by some process of selection. When thinking about this idea, I find it helpful to approach it in the following order. Hold fixed the fact that the consumers will act in a particular manner as a result of the form of each icon they receive – that they will apply a particular "interpretation rule." Suppose that *given* the rule that the consumers are following, the icon will only affect the consumers in a way which leads to them successfully performing their functions in a historically normal way if the world is such-and-such a way. Then the icon says that the world is that way (see Figure 2). To put things very roughly, the condition under which an icon is *true* is the condition under which the behaviors that the consumers will be induced to perform will be *successful*.

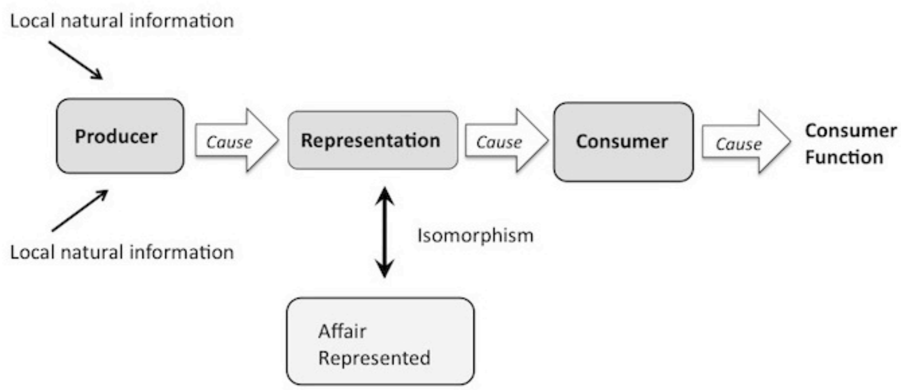

Figure 2: Millikan's schema for "descriptive" representations

Putting things less roughly and in Millkan's preferred language, icons are "supposed to map" the world, and the mapping is one that involves "transformations" of both icon and world. For example, a bee dance says there is nectar at location *X* at time *t*. The fact that this is the dance's content depends on the fact that there are *other* bee dances that could be produced at different times and with different "shapes," which have different success-conditions. If the dance had been performed at time *t*+1, it would have the content that there is nectar at *X* at *t*+1, and it would only be useful for the consumers to act on this dance by its usual rule if there was nectar at *X* at *t*+1. It is this rule mapping transformations of icons to transformations of "world affairs" that natural selection (or some similar process) singles out. [11] What selection directly acts on is the way that producers produce icons and the way consumers interpret them. Most of the possible combinations of ways of producing and consuming icons are pointless or worse. A few ways, however, are adaptive. This stabilizes the pattern of icon production and consumption. When this stabilization has occurred (according to Millikan) there will be a systematic relation visible between different forms the icon can take and different ways the world can be. That relation between "transformations" is what determines the content of a particular icon.

---

[11] The role of this talk of "mapping" and allied talk of "picturing" is discussed in detail in Shea (this volume).

As the danger-signal cases make clear, for an icon to have one of these historically normal success conditions it is not necessary that the icon's presence makes that condition probable. A low-reliability danger-signal can still be useful, and can still be a *danger*-signal rather than a signal whose content is determined by the most common cause of false alarms. Millikan's approach deals with the possibility of false content very easily. When an icon is produced and the world is not in the historically normal success condition described above, then the icon says something which is not true. This might happen 99% of the time.

The way Millikan's theory handles the possibility of false content seems an advantage over the information-theoretic approach. In some cases, though, the way Millikan's approach to content focuses on the states of the world that *explain* stabilization is at odds with reasonable-looking intuitions about content. Modifying an argument due to Pietroski (1992), suppose a producer observes the state of some variable *X*, and sends a sign that carries information about it. The consumer acts, and this action is coordinated with the state of *X*. The producer and consumer benefit – but not because of anything about the practical importance of *X*. They do well because of a correlation between *X* and another variable, *Y*, whose state cannot ever be observed by this producer and consumer but whose state has important practical consequences. And the correlation between *X* and *Y* is not due to their being at different places in a causal chain (*Y* affects *X* which affects the producer), but due to their being effects of a common cause (*Z* affects both *X* and *Y*). For Millikan, as the state of *Y* is the one that explains the stabilization of the producing and consuming, *Y* is what the signals sent by the sender are about. This can be true no matter how remote *Y* is from the producer's ability to directly track the world.

For Millikan, the content of an icon in a stabilized producer-consumer configuration is always determined by the relation between the icon and the state of the world that was causally important in the selection process that achieved the stabilization. Sometimes this fits with how we usually think of content, and sometimes it does not. Millikan can reply that she does not care about how we usually think of content – fitting with semantic intuitions is not her goal. That is fine, except that Millikan quite often defends her view by appealing to its ability to "focus" the semantic involvement of an icon on a particular part of the world, and to handle the problem of error in an intuitive way. In both those cases a defender of an information-theoretic view is also at liberty to say that he does not care about "familiar intuitions" about content.

Both views make use of real features of sender-receiver configurations in their explanation of content, and both schemes of interpretation are partly in accord with familiar intuitions about content and partly at odds with them. The features they draw on have different roles within a complete description of what sender-receiver configurations are and how they work. Millikan connects the content of a sign to the history of stabilization of the sender-receiver configuration; Skyrms links it only to the operation of the configuration at a given time. The information-theoretic approach links content to *how a sign manages to inform* a receiver; the teleo-functional view links it to what the sign is *supposed to do*, and how the world must be if the sign is to be able to do its job.

4. States of the Mind and Brain

The sender-receiver approach can be seen as a way of explaining just some semantic phenomena, or all of them. It might seem an unlikely way to approach the explanation of the content of thought. Senders and receivers are found in situations where there is *communication*, and the semantic phenomena involved in thought seem to be of a different kind. If we think about the "directedness" of thought on the world seen when someone believes that (say) that New York is south of Boston, this does not seem to be a matter of one entity *telling* something to another.

One response to this is to say that surely the brain can be seen as a signaling device. Neurotransmitters transmit signals between neurons, for a start. But whether this kind of activity fits into the sender-receiver configuration discussed above is not so clear. If we look inside a brain and find a huge network of neurons, each affected by some and affecting others, it appears that any one neuron's firing might be described as either a signal, or the reception of a signal by a receiver, or the sending of another signal, depending on how one divides things up. That is not how things are in the clear cases of sender-receiver configurations.

There is a different way of making a connection between internal states and the sender-receiver model, however. The first move is one that many philosophers will accept: endorsement of a "representational theory of the mind." On this view, our minds contain internal representations that have both semantic content and a role in cognitive processing. In virtue of using these representations in a certain way, we have propositional attitudes like the belief that New York is south of Boston. The strong and commonly held version of this idea is not just that beliefs exist as a consequence of the use of internal representations, but these inner representations have the same kind of content as that ascribed to the whole agent. Very roughly, because a representation with the content that *p* is in the right place inside the agent, the agent believes that *p*.[12] The match between what an inner sign says and what the agent believes need not be exact, and a representationalist of this kind can allow that many ordinary belief ascriptions are of "implicit" beliefs, whose contents follow from a smaller set which are "explicitly" represented. But the aim is to explain the content of thought in terms of the cognitive role of inner signs with closely related contents.

Suppose a representationalist view of this kind is provisionally accepted. Such views can be developed in several ways. Recent work on this topic has often been influenced by ideas from computer science. But there is also an older tradition. A number of philosophers have been attracted to an analogy between beliefs and maps; Frank Ramsey spoke of beliefs as maps "by which we steer" (Ramsey 1929).[13] Looking further back (as Millikan 1986 does), Plato used an analogy between knowledge and impressions on a wax tablet. So one way to have a representationalist view is to hold that inner representations are inner pictures of what they represent. Millikan thinks that this idea is still quite viable, provided that we have a suitably abstract sense of "picture," one in which spatial relations between elements of a picture need not be what is important, for example. Whether or not the idea of picturing is used, it seems possible to argue that thinking involves an interaction between parts or devices in the brain, with some producing or inscribing representations and others consuming or reading them. The language of "sender and receiver" is not especially natural here, but the model itself is applicable. Randy Gallistel has recently emphasized the importance of "read-write" memory systems in the explanation of intelligence (Gallistel 2006, Gallistel and King

---

[12] See Fodor (1981) and Field (1986).
[13] Armstrong (1973), Dretske (1988), Jackson and Braddon-Mitchell (1996).

2009). Gallistel's read-write (or write-read) architecture is another variant on the sender-receiver picture. Another link to the idea of signaling can be made by noting that *memory* might be thought of as the sending of messages through time. Many of the phenomena philosophers discuss in terms of "belief" can be understood as the operation of memory.

So a possibility has been sketched, but what would these inner senders and receivers (writers and readers, producers and consumers) have to be like? One way to put the point is to ask whether there are things we might discover that would show that the sender-receiver model was *not* applicable to thought, and if we discovered one of those things, what that would show about beliefs.

I will compare how these issues are handled in some different parts of Millikan's work. In her 1984 book Millikan discusses, with some caution, the application of her analysis of intentional icons to beliefs (and similar states) in human beings. She outlines a way in which the analysis *could* apply, a way which depends on the existence of a language-like medium for thought and some specific mechanisms which interact with inner sentences. The "interpreter" of beliefs she hypothesizes is a device called a "consistency tester" (1984, pp. 144-46). A consistency tester looks for ways in which contradictions (which are to be avoided) may arise in a single mind, by comparing different ways in which inner sentences can be formed.

If "consistency testers" are real parts of us, they are presumably very contingent parts of the machinery. In Millikan's discussion, their usefulness derives from the fact that we use inner sentences that contain newly coined mental "terms." It seems that humans could get by fairly well without consistency testers, even if we think in a language-like medium, though we might produce more contradictory combinations of inner sentences without them. It seems unlikely that our having beliefs depends on their presence.

In the same discussion Millikan also describes some fictional analogues of humans, whose concepts are all innate and products of evolution by natural selection. These fictional agents are said to contain some components that produce inner sentences, and other components that use the sentences in guiding behavior (1984, Chapter 8). Millikan takes this to be unproblematic in principle, and so do I; it is a possible psychology. So in Millikan's 1984 book we have inner sentences that qualify as icons within a fiction, and empirical hypotheses about a "consistent tester" which, if real, would make human beliefs into intentional icons as well. What is missing is something in between the acknowledged fiction and the very strong hypothesis about consistency testers.

Millikan's follow-up work about the mind (1986, 1989, 2004) gives a different treatment of interpreter or consumer mechanisms. Consistency testers are not made central, and neither is a language-like medium for thought. In some of these discussions, the notion of a consumer is apparently understood in a very broad way; all the machinery guiding action and inference that is affected by a representational state is a "consumer" of that state. In "Biosemantics" (1989) Millikan says: "Let us view the system, then, as divided into two parts or two aspects, one of which produces representations for the other to consume" (p. 295). I am not sure whether this is meant as an idealization which will help us think about the problem, a commitment to something like a "belief-box" model of thought, or as a minimal commitment that must surely be accurate in virtue of uncontentious facts about the causal flow in the brain from senses to effectors.

The last of these options seems unlikely; a "consumer" should not just be everything downstream of an inner state. And, given that, we should think about the possibility that what evolution has done in the case of brains is move away from the producer-consumer configuration. Perhaps, for instance, it has built something in which the representation role "leaks into" that of producer or consumer. The configuration may also be present in a partial or attenuated form.

I will follow this idea up by looking at some work in psychology, useful here because it seems to be a promising case for both representationalism and Millikan's producer-consumer model. Earlier I mentioned philosophers' description of beliefs as "maps by which we steer." Psychologists also talk of inner maps, especially when studying navigation. This work has been discussed rarely by philosophers, but it provides a useful alternative empirical angle on the idea of internal representation, distinct from "language of thought" views influenced by linguistics and classical AI.[14]

---

[14] Exceptions include Bermudez (2003), Rescorla (2009), and a short discussion in Millikan (2004).

Psychological work on inner maps originates mainly with E. C. Tolman, especially in his "Cognitive Maps in Rats and Men," (1948). This paper was part of an attempt to show that "stimulus-response" models are inadequate for both kinds of animal. In Tolman's discussions, evidence for an inner map is behavioral. Experiments show that rats can be smarter in dealing with space than a stimulus-response view allows. For example, in a "starburst maze" experiment, a rat first learns a highly indirect route to a food source, and is then presented with a range of novel paths which lead more or less directly to the food. Rats choose a nearly direct path much more often than chance would predict. This, for Tolman, shows that the rat stores information about its environment in a map-like representation which can later be used to work out how to behave.

Some years after Tolman, evidence was found that whatever the spatially smart rats were doing, the hippocampus was the likely part of the brain where it happened. The hippocampus contains "place cells," which fire when the animal is in a particular location, regardless of the animal's orientation (O'Keefe and Nadel 1978). Neurobiological work has continued. A feature of this work, as I understand it, is that a picture has *not* emerged in which there is a separation between an inner map and a "reader" of the map. Maybe such a picture will emerge, but it has not so far. One alternative is that the important neural structures have map-like and reader-like features mixed together. "Place cells," for example, are often described informally in this mixed sort of way – as representations of where the animal is, which may be read, and as entities which have done some reading and know where they are.

I will look closer at the issue with the aid of a computational model of how a system of mapping of space might work in the brain, a model due to Reid and Staddon (1997, 1998). Their aim is not to model how the map is formed, but how a map can be used by the rat, once it exists.

The model assumes a connected two-dimensional lattice of units. These units are a bit like neurons, or might better be seen as small assemblies of neurons. Each unit is connected to eight neighboring units, its N-S-E-W neighbors plus diagonals. Each unit is also connected to sensory input and motor output. At any time, just one unit in the lattice is *active*, and the others are *inactive*. For each unit there is a unique sensory stimulus which makes it active. This amounts to the animal's ability to track its present location in the environment. Adjacent units in the lattice are cued by their stimuli to adjacent locations in space.

Each unit in the lattice has at each time a value of a variable *V*. As time passes, the value of *V* for all the inactive units changes by a kind of "diffusion." Each unit is affected by its neighbors in a gradual ongoing averaging process. If a low-value unit is between a lot of high-value ones, it will approach their higher value and they will be slightly affected by its lower value. There is an exception to the diffusion rule. If two neighboring units correspond to locations that are separated by a barrier that the animal has found that it cannot move through, there is no diffusion between those units.

The *active* unit at a given time has its value of *V* set differently. The animal is either rewarded or not rewarded at each time step. If the animal is rewarded, the *V* value of its active unit is set to a high value, and it is set to a low value if the animal is not rewarded.

Lastly, there is a rule of movement for the animal. It moves at each time step in a direction determined by the neighbor with the highest value of *V* at that time step (with the proviso that movement cannot occur across a barrier). This is an entirely local process; the active unit at a time step sends a motor output determined by the *V* values of all its neighbors, and by nothing else. No inactive units emit motor outputs. If the active unit has a higher *V* than its neighbors, the animal stays put. So, for example, if at time *t* the active unit has a neighbor in the lattice to its north-east which has a higher *V* value than all the other neighbors and itself, the animal moves in a direction which makes that north-east neighbor the new active unit.

Reid and Staddon show how this simple process suffices to generate some quite smart navigation behaviors, including short-cut and detour behaviors, and solution of some maze problems (including the "starburst maze"). They show this with simulations in some cases and verbal arguments in others. The model assumes than an animal goes through an initial exploration or training period, and in a "trial" phase is released and allowed to look for food. The feature of the mechanism which generates apparent spatial "insight" is as follows. If an animal learns the location of food, by being present at that place during the training period, the corresponding unit will acquire a high *V*, which will then diffuse out from the unit corresponding to that location after training has ceased.

This generates a "landscape" of *V* values which is affected by encountered barriers as well as the distance and direction of food. In the trial phase, the animal moves by hill-climbing from lower to higher values of *V*, leading it from a starting location to food. "Thus, Tolman's contention that something more than *S-R* [Stimulus-Response] principles is required to account for maze behavior is only partly true. A map *is* required (in our scheme), but beyond that, no "insight," no internal cyclopean eye with an overview of the situation as a whole, is necessary" (Reid and Staddon 1997, p. 227)

Suppose that an evolved system did deal with space in the way described by Reid and Staddon. How would this relate to a scheme of interpretation for inner states based on a producer-consumer model like Millikan's?

The state of the lattice and its relation to the rest of the animal embodies the animal's knowledge of the spatial structure of its environment. That seems true. It is also true that in one sense, the lattice is a map of the landscape. *You* could look at the lattice in the brain and use it as a map; you could look at the lattice and read off the pattern of barriers, because the channels through which diffusion occurs are presently blocked or broken there. But, I will argue, there is no reader of the map, separable from the map itself, inside the animal.

There may be *a* kind of reading going on, at a low level. At each time-step, one unit is active. The active unit compares the *V* values of neighboring cells, and responds to these values by choosing a direction of movement. With the sensory stimulus that ensues, a new unit then becomes active. The new active unit then reads the *V* values of its own neighbors, leading to a new motor output.

I said just then that the active unit "reads" the *V*'s of its neighboring cells. That way of putting things might be denied – certainly it is not a clear case of the producer-consumer configuration. It is hard to say what the "producer" is, as the entire lattice and its history of reinforcement and exploration is responsible for the *V* values found at a unit at a time. Perhaps it is better to say that the units form an interacting network which achieves a kind of computation, but no sending and receiving of messages. (Or maybe messages exist only at an even lower level, where a neurotransmitter is an icon-like intermediary telling unit *i* what the present *V* value in unit *j* is.) But rather than making this a sticking-point, let's instead accept that there is a low-level reading of signs going on when an active unit tracks and responds to the values of *V* around it. Then the *V* values that are read by the active unit could be interpreted as having content involving environmental conditions. We might initially say that each *V* value that is read maps *the closeness of reward along a viable path*, at that time. That would not be right, though, as a given *V* value has no significance on its own, and what matters is the comparisons between the *V*'s of the neighbors. The overall level of *V* across the landscape becomes more similar as the animal moves further in time from its most recent reward. So it would be more accurate to say that the total set of *V* values of an active unit's neighbors maps the direction of a viable path to reward. Employing Millikan's schema, the active unit is supposed to respond to these *V* values in a way that will lead to the system doing well only if such-and-such *is* the direction of a viable path to reward.

So at this unit-to-unit level, a description in terms of the consumption of signs might be given. But the reading event that occurs at each time-step is sensitive only to the *V* values (and any barriers) in the immediate neighborhood of the active unit. Whenever the animal moves, a new unit becomes active. What was *part of the map* at step *t* – an element important only in virtue of its readable *V* value and sensitivity to the diffusion process – becomes *the reading device*, the consumer of information, at step *t*+1.

Reid and Staddon call one of their papers "A Reader for the Cognitive Map" (1997). But what they have described is more like a self-reading map, or a map which works without being read at all. The total set of *V* values in the lattice at a time is a map of where rewards are found the landscape according to the animal's experience. Again, we as outsiders could look at the lattice and see not just where the barriers are, but also how to jump to food (find the high point where *V* is diffusing away from). But the animal can't do this. Only the presently active unit treats the local *V* levels as informative. That unit instructs a small step in some direction, and that unit then falls back into being part of the map as a new unit becomes active. The new active unit, which will be one of the former active unit's neighbors, will read the former active unit in turn. So there is lots of low-level *reading of V-values*, by readers which are separable from the things being read, but there is not in the same sense a *reader of the map*.

Reid and Staddon's is a "how-possibly" model. It illustrates one way that things might turn out in the explanation of how brains work, and this includes interesting possibilities with respect to the relation between personal and sub-personal levels. Looking at the rat and its behavior, we initially think: there is a map of the landscape in there, which has a reader, and that is how the animal knows where food is. Then we imagine finding a Reid-Staddon mechanism inside the rat. We find there is a low level at which a producer-consumer structure can be seen, though of an unclear kind. The contents of the inner icons we recognize there *might* also be seen as corresponding to belief-like states of the whole animal – as things the animal thinks or knows. But the organism-level contents we thought we were analyzing – knowledge of the layout of the environment, knowing where the food is – are not being found. There are no icons in the system with contents like that. The Reid-Staddon model posits a map that *would* be readable, if a suitable reader was present. But the rat cannot read the map in the way we can. (In this respect the Reid-Staddon map is like the "somatosensory map" in the human cortex.)

On seeing such a mechanism, we might decide it was a mistake to say the animal knows the layout of the landscape, and knows where the food is. The animal does not know anything like that, we might decide, but only knows at each moment the direction of the best path (which may be different from the direction of food). We might instead decide that our initial attribution was not *wrong*, but move to a more instrumentalist or behaviorist treatment of attributions of these belief-like states. Or we might say that in virtue of all the facts about how various sub-personal signal systems work, along with much else, the rat really does have the whole-organism beliefs we initially attributed. Then the producer-consumer model will be telling us about sub-personal contents that provide a *basis* for the whole organism's intentional states, even though the contents of those states are not the contents of any inner signs.

The picture illustrated in a simple and mild form here might turn out to apply more generally. This is a picture of cognition in which there is much low-level natural signaling – perhaps with leakage of sign into consumer, perhaps with unclear producers – but with no straightforward relation between sub-personal sign contents and the agent's intentional states. Or maybe that is generally true, but with an important exception: cases where the inner sign is an internalized public language sentence which functions in cognitive processing.[15] In those cases, an inner sign does become a vehicle by which an agent can have thoughts with a particular sophisticated content, and where the sign has the same content as the thought.

An analogy might be drawn with another controversial realization of the sender-receiver configuration, the case of genes. Protein synthesis within a cell is a process in which a kind of "reading" of the genes does go on, along with a quasi-logical regulation of gene action by regulatory mechanisms. This, again, is a questionable case from the point of view of section 2, but let's accept it here. It is also true that a lot of highly coordinated gene-reading of this kind, taking place within many different cells, results in the growth and development of an entire organism. But there is no reader of a "genetic message" at the whole-organism level. If there is a kind of content in the messages being read, it is restricted to the specification of the primary structure of protein molecules. Nothing is reading a specification of, or instructions to make, a hand or a kidney. We could have had a Grand Central Reader of the genome but we do not.

The sender-receiver configuration is a "natural kind," one that Millikan has done a great deal to help us understand. This kind is found in a variety of forms, arising independently many times. We have some account of how it comes to exist and is maintained. Our investigation of these set-ups is also affected, perhaps both positively and negatively, by habits of description and interpretation that derive from the place that sign systems have within everyday life. Among the many further questions that might be considered are those about when the clear, paradigmatic forms of the configuration can be expected to arise, and when instead the attenuated and blurred ones will, as a consequence of the role being played by the set-up, its circumstances, and the raw materials being fashioned into its various components.

Acknowledgment: I am grateful to Rosa Cao, Ruth Millikan, Nick Shea, and Brian Skyrms for discussions and correspondence.

[15] See Dennett (1991), Caruthers (2002), Yegnashankaran (forthcoming).

References

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

Bergstrom, C. and M. Rosvall (forthcoming). The transmission sense of information. *Biology and Philosophy*.

Bermudez, J. L. (2003). *Thinking Without Words*. New York: Oxford University Press.

Camilli, A. and and B. L. Bassler (2006). Bacterial small-molecule signaling pathways Science 311: 1113 - 1116.

Caruthers, P. (2002). The cognitive functions of language. *Behavioral and Brain Sciences* 25: 657-674

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

Dretske, F. (1988). *Explaining Behavior*. Cambridge MA: MIT Press.

Dennett, D. C. (1991). *Consciousness Explained*. Boston: Little, Brown, and Co.

Eco, U. (1976). *A Theory of Semiotics*. Bloomington: Indianapolis University Press.

Field, H. (1986). Mental representation. *Erkenntnis* 13: 9-61

Fodor, J. (1981). *Representations*. Cambridge MA: MIT Press.

Fodor, J. A. (1984). Semantics, Wisconsin style. *Synthese* 59: 231-250

Gallistel, C.R. (2006). The nature of learning and the functional architecture of the brain, in Q. Jing, et al (eds) *Psychological Science Around the World, Vol.1. Proceedings of the 28th International Congress of Psychology*. Psychology Press, Sussex. pp. 63-71.

Gallistel, C.R., and A. King (2009). *Memory and the Computational Brain: Why Cognitive Science will Transform Neuroscience*. New York: Blackwell/Wiley

Godfrey-Smith, P. (1989) Misinformation. *Canadian Journal of Philosophy* 19: 533-550.

Godfrey-Smith, P. (1992). Information and adaptation. *Synthese* 92: 283-312.

Godfrey-Smith, P. (1996). *Complexity and the Function of Mind in Nature*. Cambridge: Cambridge University Press.

Haig, D. (1996) Placental hormones, genomic imprinting, and maternal–fetal communication. *Journal of Evolutionary Biology* 9: 357-380.

Haig, D. (2008) Conflicting messages: genomic imprinting and internal communication. Pages 209-223, in P. D'Ettorre and D. P. Hughes. *Sociobiology of communication*. Oxford University Press, Oxford.

Haig, D. (2010) Transfers and transitions: parent-offspring conflict, genomic imprinting, and the evolution of human life history. *Proceedings of the National Academy of Sciences, USA* 107: 1731-1735.

Hamilton, W. D. (1975). Innate Social Aptitudes of Man: An Approach from Evolutionary Genetics. In R. Fox (ed.) *Biosocial Anthropology*. London: Malaby Press, pp. 133 –155.

Holldobler. B. and E. O. Wilson (2008). *The Superorganism: The Beauty, Elegance and Strangeness of Insect Societies*. New York: W. W. Norton.

Jackson, F. and D. Braddon Mitchell (1996). *Philosophy of Mind and Cognition*. London: Blackwell.

Lachmann, M., S. Szamado, and C. T. Bergstrom (2001). Cost and conflict in animal signals and human language. *Proceedings of the National Academy of Sciences USA*. 98:13189-13194

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

Maynard Smith, J. (2000). The concept of information in biology. *Philosophy of Science* 67: 177—194

Maynard Smith, J. and D. Harper (2003). *Animal Signals*. Oxford: Oxford University Press.

Millikan, R. G. (1984). *Language, Thought, and Other Biological Categories*. Cambridge MA: MIT Press.

Millikan, R. G. (1986) Thoughts without laws: Cognitive science with content. *Philosophical Review* 95: 47-80

Millikan, R. G. (1989). Biosemantics. *Journal of Philosophy* 86, 281–97.

Millikan, R. G. (2004). *Varieties of Meaning*. Cambridge MA: MIT Press.

O'Keefe, J. and L. Nadel (1978). *The Hippocampus as a Cognitive Map*. Oxford: Clarendon Press.

Pietroski, P. (1992). Intentionality and teleological error. *Pacific Philosophical Quarterly* 73: 267-81.

Ramsey, F. P. (1929). "General Propositions and Causality" in his *Philosophical Papers*, ed. D. H. Mellor. Cambridge: Cambridge University Press, 1990, pp. 145–63.

Reid, A. and J. Staddon (1997). A reader for the cognitive map. *Information Sciences*, 100: 217-228.

Reid, A. and J. Staddon (1998). A dynamic route-finder for the cognitive map. *Psychological Review* 105: 585-601

Rendall, D., M. Owren, and M. Ryan (2009) What do animal signals mean? *Animal Behaviour* 78, 233–240.

Rescorla, M. (2009). Cognitive maps and the language of thought. *British Journal for the Philosophy of Science* 60: 377-407

Searle, J. (1992). *The Rediscovery of the Mind.* Cambridge, MA: M.I.T. Press.

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

Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal* 27: 379 – 423, 623 – 56.

Shea, N. (2007). Representation in the genome and in other inheritance systems. *Biology and Philosophy* 22: 313-331.

Shea, N. (this volume). Millikan's isomorphism requirement.

Skyrms, B. (1995). *Evolution of the Social Contract*. Cambridge: Cambridge University Press.

Skyrms, B. (2010). *Signals*. Cambridge: Cambridge University Press.

Stegmann, U. (2009). A consumer-based teleosemantics for animal signals. *Philosophy of Science* 76: 864–875

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review* 55: 189-208

White, J. (1979). The plant as a metapopulation. *Annual Review of Ecology and Systematics* 10: 109 – 45.

Yegnashankaran, K. (forthcoming). Reasoning as action.