Propositional content in signalling systems

Jonathan Birch

Christ's College, University of Cambridge

jgb37@cam.ac.uk

(forthcoming in *Philosophical Studies*; the final publication will be available at link.springer.com)

Abstract: Skyrms, building on the work of Dretske, has recently developed a novel information-theoretic account of propositional content in simple signalling systems. Information-theoretic accounts of content traditionally struggle to accommodate the possibility of misrepresentation, and I show that Skyrms's account is no exception. I proceed to argue, however, that a modified version of Skyrms's account can overcome this problem. On my proposed account, the propositional content of a signal is determined not by the information that it *actually* carries, but by the information that it *would* carry at the nearest separating equilibrium of the underlying evolutionary dynamics. I show that this amended account yields reasonable ascriptions of false propositional content in a well-known formal model of the evolution of communication (the 'Philip Sidney' game), and close with a discussion of the serious but perhaps not insuperable difficulties we face in applying the account to examples of signalling in the real world.

1. Signals, information and propositions: the state of play

Many social animals have complex signalling systems that help them respond adaptively to the threat of predation. Vervet monkeys (*Chlorocebus pygerythrus*) are particularly famous for this: on observing a predator, a vervet varies the alarm call it emits depending on whether the predator is an eagle, a leopard or a snake, thereby allowing other nearby vervets to take appropriate evasive action (Seyfarth et al. 1980; Cheney and Seyfarth 1990). Comparably sophisticated alarm call systems have since been discovered in numerous Old and New World monkeys (Zuberbühler 2000; Di Bitetti 2003), in lemurs (Macedonia 1990), in chimpanzees (Slocombe and Zuberbühler 2005), in meerkats (Manser 2001, 2009; Manser et al. 2001), and in at least six avian species (Evans and Evans 2007; Gill and Bierema 2013).

There is a strong intuitive temptation to ascribe propositional content to these alarm calls. Faced with such impressively sophisticated signalling behaviour, we naturally want to say that each alarm call represents a particular state of affairs: that, for instance, a certain type of alarm call represents the proposition *that an eagle is nearby*; whereas a different type of alarm call, with a different set of acoustic characteristics, represents the proposition *that a leopard is nearby*. Many cognitive ethologists share this intuition: the predator-specific alarm calls of vervet monkeys and other species are routinely described as 'semantic' or 'representational' or 'referential' or 'functionally referential' in the ethological literature (see above citations; see also Evans 1997; Townsend and Manser 2013 for reviews).

Yet, irresistible as they often seem, ascriptions of propositional content to animal signals are deeply controversial in both ethology and its philosophy. Everyone can agree that a particular type of vervet alarm call *co-varies*, to some extent, with the presence or absence of an eagle. Consequently, it is fairly (though not wholly[1]) uncontroversial that there is *mutual information* between the alarm call and the presence or absence of an eagle, in the technical, mathematical sense in which that term is employed in information theory. In other words, knowing whether or not the alarm call is sent *reduces our uncertainty* about whether or not there is an eagle nearby (Shannon 1948; Shannon and Weaver 1949). But I suspect that many philosophers—and some ethologists too—would want to stop here. They would be reluctant to take the further step of saying that animal signals carry genuine propositional content: that they represent external states of affairs in the same full-blooded, semantic sense as the sentences of a human language.[2]

The impression of a serious, perhaps unbridgeable gap between information-sensu-Shannon and genuine propositional content arises chiefly from the notorious 'problem of

[1] See Rendall, Owren and Ryan 2009; Owren, Rendall and Ryan 2010; Rendall and Owren 2013 for criticism of even this seemingly innocuous form of informational terminology in the context of animal communication.

[2] See Adams and Beighley 2013 for a particularly clear statement of this sceptical view.

error' (Fodor 1984; Godfrey-Smith 1989; Crane 2003). Propositions have truth-values and truth conditions: they can be true or false, depending on whether or not their truth conditions obtain. The upshot is that the sentences of a human language don't have to represent the world truthfully: we can, if we want to, use language to express falsehoods. The possibility of expressing falsehoods—of *mis*representation—goes hand in hand with the notion that one's utterances have propositional content. Yet information theory seems unable to explain how misrepresentation is possible. However one cashes out the details, a signal can be said to 'carry the information that *s* is *F*' in the sense of information theory only if it stands in a certain natural relation—roughly, a relation of *reliable indication*—to *s*'s being *F* (see especially Dretske 1981). Hence, if a signal carries the information that *s* is *F*, it follows that *s* is *F*. Nothing in the theory allows us to makes sense of how a signal could still carry the information that *s* is *F* even though *s* is not *F*.

For a concrete example, suppose a vervet makes its characteristic leopard alarm call on observing a cheetah.[3] If, as intuition would have it, this type of alarm call represents the proposition that a leopard is nearby, then its propositional content is false on this occasion. But note that the alarm call will statistically co-vary with the presence or absence of a cheetah in much the same way that it co-varies with the presence or absence of a leopard. Because of this, it seems that information theory cannot underwrite the intuitive asymmetry between the two cases. At most, it allows us to say that, in both cases, the signal carries the true disjunctive content that there is a leopard *or* a cheetah nearby.[4] But it does not explain how the signal could falsely represent the presence of a leopard.

---

[3] This does in fact happen from time to time (see Cheney and Seyfarth 1990, p. 169). Crane (2003) also uses this example to illustrate the problem of error.

[4] In a sense, the 'problem of error' and the 'disjunction problem' are two sides of the same coin: information theory struggles to explain misrepresentation *because*, in cases where a signal can be caused by any of a number of factors, it appears to imply that the signal must truthfully represent the disjunction of these factors. But a number of authors have argued that the disjunction problem is more pervasive than the problem of error, and I need not take sides on this issue here (though see footnote

This is well-trodden ground: there are many putative solutions to the problem of error, and many objections and counterexamples to those solutions (landmarks include Stampe 1977; Dretske 1981, 1988; Fodor 1984, 1987, 1990; see Adams and Aizawa 2010 for a review). I do not intend to survey these here, though I will mention some in due course. It is fair to say, however, that the past three decades have not produced any solution that is widely accepted or unproblematic. Indeed, a pessimistic thought has increasingly taken hold: the thought that no adequate solution to the problem of error is possible (see Godfrey-Smith 1989; Crane 2003). If that is right, then the gap between information-sensu-Shannon and genuine propositional content is truly unbridgeable. Of course, we might still be tempted to give an *informal* propositional gloss to the information a signal carries, but it would be a mistake to take that gloss too seriously. It would be a mistake to think that we could ever use information theory as the basis for *serious* attributions of propositional content to animal signals, or indeed to anything else.

In this article, I want to suggest a new way out of what appears a bleak predicament for information-theoretic accounts of content. Though I will focus on content in simple signalling systems—of the sort we find among vervet monkeys, lemurs and meerkats—my hope is that the solution I develop will extend to more complex cases too. The outline of the article is as follows. In Section 2, I explain and scrutinize a novel information-theoretic approach to content recently developed by Brian Skyrms (2010a). Skyrms suggests that the propositional content of a signal is straightforwardly explicable within the formal apparatus of information theory, because it is nothing more than a special case of a more general notion of informational content. In Section 3, I show that Skyrms has not evaded the problem of error. Yet, despite the failure of his specific proposal, Skyrms's work still points in the direction of a promising approach to the problem, and I develop this approach in Section 4.

---

22). For further discussion of the relationship between the two problems, see Fodor 1990; Neander 1995, 2012; Crane 2003; Adams and Aizawa 2010.

The basic line of thought I develop is this. Skyrms's overarching project in *Signals* can be viewed as an attempt to integrate information theory with evolutionary game theory; and an integration of these theories has extremely helpful consequences for naturalistic theories of content. Why? Because, several decades ago, Dennis Stampe suggested (in broad terms) that the key to formulating a successful causal/informational theory of content is to specify relevant *fidelity conditions*, such that the content of a representation is determined by the state of affairs that it *would* reliably indicate if the fidelity conditions were satisfied (see Stampe 1977, p. 51). The challenge was to specify the relevant fidelity conditions in a sufficiently precise and principled way.

My suggestion (in short) is that Skyrms, by explicitly connecting information theory to evolutionary game dynamics, allows us to do that. For once this connection is in place, we can meaningfully ask (and determinately answer) questions such as: what *would* the informational content of these signals be, if the underlying dynamics satisfied such-and-such conditions, or if the population were at an equilibrium with such-and-such properties? This opens up the possibility that the propositional content of a signal might be determined not by the information it *actually* carries, but rather by the information it *would* carry if relevant fidelity conditions, formally articulated in evolutionary-game-theoretic terms, were satisfied. In Section 4, I propose a new information-theoretic account of propositional content along broadly these lines. In Section 5, I show (by means of a concrete example) how it handles the problem of error. In Section 6, I consider some problems the account still faces. Section 7 draws together the discussion and concludes.

One final preliminary remark is in order, so as to block a natural concern about the overall project. If the ultimate goal here is an account of how propositional content arises from informational connections between representations and the world, why start with *signals*? Should we not start (following Grice 1957) with *mental* representations, and then explicate the meaning of signals (following Lewis 1969) in terms of the intentions of the sender and the beliefs of the receiver? This has been the usual way of proceeding for a number of decades now, but I am not convinced it is the best way. Mental representation is, after all, no less mysterious a phenomenon than signalling. Indeed, if we could come to understand the intentionality of signals *without* presupposing any capacity for mental representation on the part of the sender or receiver, we could contemplate the possibility of turning the traditional explanatory relationship on its head, and explaining mental representation as a form of internalized signalling. I will not pursue this project here, but it is one important underlying motivation for a 'signals-first' approach to content.

2. Skyrms on the content of signals

Skyrms's work on signalling, in his (2010a) book *Signals* and in associated papers (Skyrms 2008, 2009, 2010b, 2012; Skyrms and Huttegger 2013) is a systematic attempt to build on foundations laid by David Lewis (1969) and Fred Dretske (1981). Lewis (1969) introduces the idea of a *signalling game*: a simple, formal model of communication involving two agents, a sender and a receiver. In a signalling game, the sender can vary the signal she sends depending on the state of the world she observes, and the receiver can vary her actions depending on the signal received. Lewis argues that conventional signalling systems, in which particular signals are taken by both agents to represent particular states of the world, emerge at the coordination equilibria of these games.[5] We see, in other words, the origin of propositional, proto-linguistic meaning as a conventional solution to a coordination problem.

Skyrms revises and augments Lewis's framework in several important respects. First, he moves from *classical* to *evolutionary* game theory: the convention-establishing work that is done by conscious rational choice in Lewis's models is done by natural selection or trial-and-error learning in those of Skyrms. If the models are interpreted as models of evolution by

[5] A coordination equilibrium is a special kind of Nash equilibrium. At a Nash equilibrium, no agent gains by unilaterally switching from one strategy to another. At a coordination equilibrium, no agent gains if *anyone* unilaterally switches.

natural selection, then the agents—the senders and receivers—need not be capable of rational choice in any sense, allowing in principle for the inversion of the traditional Gricean strategy: we need not regard signalling as in any way dependent on pre-existing mental representation. Second, Skyrms moves from purely static analyses based on the concept of a coordination equilibrium to dynamic modelling of how signalling systems evolve over time. Third, and most relevantly for our purposes, Skyrms combines his game-theoretic models with an information-theoretic account of what it is for a signal to carry *informational content*: an account that can be seen as a generalization of that of Dretske (1981). Skyrms proceeds to argue for a very close relationship between a signal's informational content and its propositional content: a relationship so close, in fact, that the former straightforwardly determines the latter.

2.1 Informational content

The thought behind Skyrms's account is that the informational content of a signal should depend on how the objective probabilities of states of the world change when we conditionalize on the signal being sent. From the outset, Skyrms helps himself to an objective notion of probability; moreover, he assumes that objective probabilities can be identified with relative population frequencies (Skyrms 2010b, p. 157). One might well worry about these assumptions; for while it may be reasonable to identify objective probabilities with relative frequencies in the infinite populations with which Skyrms's models are typically concerned, this identification seems rather more dubious in finite populations. I intend to put these issues to one side here. I will simply follow Skyrms in assuming that the probabilities his models describe do admit of an objective, frequentist interpretation.

To see how these objective probabilities give rise to an objective notion of informational content, suppose we have a signalling game in which we the world can occupy of four states: S_1, S_2, S_3 and S_4. To make the example more vivid, we can suppose that these states have a concrete biological interpretation: S_1 is a nearby eagle, S_1 is a nearby leopard, S_3 is a nearby snake, and S_4 is the absence of any of the above predators. Let us also introduce the symbol σ_1 for some acoustically distinctive type of alarm call, without making any prior judgements about what this alarm call 'means'. For any particular state (S_1, say), we can compare the objective probability of that state *conditional on a token of* σ_1 *having being sent* to the state's unconditional objective probability. We can write this comparison in the form of a fraction:

$$\frac{P(S_1|\sigma_1)}{P(S_1)}$$

This fraction will be less than 1 if σ_1 lowers the probability of S_1, greater than one if it raises the probability of S_1, and exactly 1 if it leaves the probability of S_1 exactly unchanged.

The next step is to take a logarithm (conventionally, the log to the base 2) of this number:

$$\log_2\left(\frac{P(S_1|\sigma_1)}{P(S_1)}\right)$$

The effect of taking a log is that, if σ_1 lowers the probability of S_1, we now have a negative number (since logs of numbers less than 1 are negative); whereas if σ_1 raises the probability of S_1 we have a positive number. So the sign of the number tells us the direction in which the signal shifts probabilities. If σ_1 leaves the probability of S_1 exactly unchanged, the log will be zero.

If we repeat this procedure for all four states, we end up with a list of four numbers that give us an overall sense of how σ_1 shifts the probabilities of states of the world. We can write this list as a vector:

$$\mathbf{I}_{\sigma_1} = \left\langle \log_2\left(\frac{P(S_1|\sigma_1)}{P(S_1)}\right), \log_2\left(\frac{P(S_2|\sigma_1)}{P(S_2)}\right), \log_2\left(\frac{P(S_3|\sigma_1)}{P(S_3)}\right), \log_2\left(\frac{P(S_4|\sigma_1)}{P(S_4)}\right) \right\rangle$$

Skyrms's proposal is that this vector *fully characterizes* the informational content of σ_1. Informational content just *is* a vector that tells us how σ_1 shifts the probabilities of states of the world. The vector will have as many elements as there are relevant states.[6] For instance, the vector might look something like this:

$$\mathbf{I}_{\sigma_1} = \langle 1.0, 0.26, -1.32, -1.32 \rangle$$

This is the informational content vector of a signal that raises the probability of S_1, slightly raises the probability of S_2, and lowers the probabilities of S_3 and S_4 by an equal amount.

2.2 Propositional content

On the face of it, this list of numbers bears little resemblance to propositional content as we know it. There is no *that* clause, no *p* placeholder, no obvious place for a proposition to slot in. Skyrms argues, however, that propositional content is simply a special case of this broader, quantitative notion: that we can, in other words, read off the propositional content of a signal from its informational content vector.

Specifically, Skyrms proposes that we identify the propositional content of a signal with the disjunction of all possible states of the world that the informational content vector *does not rule out*—that is, the disjunction of all possible states of the world with which the informational content vector is logically compatible. For example, suppose a signal σ carries the following informational content:

$$\mathbf{I}_{\sigma} = \langle 1.26, 0.68, -\infty, -\infty \rangle$$

[6] Of course, individuating 'relevant states' can be a significant challenge outside the world of simple formal models. I return to this problem in Section 6.

This is the informational content vector of a signal that raises the probability of S_1 and S_2, while ruling out S_3 and S_4, in the sense that it shifts their probabilities all the way to zero.[7] Skyrms's suggestion is that, if a signal has this kind of informational content vector—a vector that rules out some states of the world completely—then we can attribute substantive propositional content to the signal. In this case, we can say that the signal carries the propositional content *that S_1 or S_2 obtains*. More generally, the propositional content will be the disjunction of states for which the vector does not have an entry of '$-\infty$'. For Skyrms, therefore, no deep gulf separates informational content and propositional content. Propositional content arises whenever a signal has informational content that rules out one or more states of the world.

3. Skyrms meets the problem of error

Dretske's (1981) information-theoretic approach to content had no difficulty accommodating cases in which a signal reliably and truthfully indicates the state of affairs it represents. The problem was that it struggled to explain how a signal could ever represent a state of affairs when tokened in its absence. Does Skyrms's new account fare any better? In other words, does it overcome the problem of error? The answer, regrettably, is no. For it is impossible, on Skyrms's account, for any signal to ever carry false propositional content.[8]

To see why, we just need to think about what it would take, on Skyrms's account, for a signal σ to have false propositional content. Two conditions would have to be satisfied. Firstly, it would have to be the case that the informational content vector of σ rules out a state of the world, S_x, in the sense that $P\left(S_x \middle| \sigma\right) = 0$ (call this *Condition A*). Note that, given

---

[7] Skyrms here assumes that $\log\left(0\right) = -\infty$. Strictly speaking, we should say that $-\infty$ is the limit of $\log\left(x\right)$ as $x \to 0$, since $\log\left(0\right)$ is undefined. But Skyrms's assumption seems innocuous enough in the present context.

[8] Godfrey-Smith (2013) briefly notes this problem. Here I aim to spell it out in more detail.

Skyrms's frequentist understanding of probability, this implies that σ is *never sent* when S_x obtains. Only if Condition A is satisfied can Skyrms's account attribute a propositional content that implies $\sim S_x$ to σ. Secondly, it would have to be the case that, at on at least one occasion, σ is sent when S_x obtains, so that on at least one occasion the propositional content of σ is false (call this *Condition B*). The problem is that Conditions A and B can never be satisfied simultaneously. Condition A holds only if $P(S_x|\sigma) = 0$ (i.e., the signal is *never* sent when S_x obtains); whereas Condition B holds only if $P(S_x|\sigma) \neq 0$ (i.e., the signal is sent on at least one occasion when S_x obtains). So it is impossible, on Skyrms's account, for σ to carry false propositional content. Since this argument makes no assumptions about the nature of the signal σ or of the state S_x, the conclusion is quite general: if Skyrms's account is correct, then no signal can ever carry false propositional content. [9]

On the face of it, then, Skyrms's account is no more successful than its Dretskean precursor in evading the problem of error. It gives us a richer account of the easy cases, but it fares no better in the difficult cases. Where do we go from here? One option, of course, would be to explicitly limit the scope of the account. It was only ever intended as an account of *natural* meaning, we might say; and, as Grice (1957) emphasized, 'natural meaning' should

---

[9] Plainly, the problem here stems from the fact that Skyrms's account (like Dretske's before it) implies that a signal must shift the probability of $\sim p$ to 0 in order to carry the propositional content that *p*. Given this, one might naturally suspect that a 'quick fix' is available, whereby we simply relax this requirement and attribute propositional content whenever the probability of $\sim p$ is downshifted to, say, 0.01. As Dretske realized from the outset, the problem of error would not arise for such an account—at least not in this acute form. However, Dretske (1981, Chs 2-4) offers two main arguments against relaxing this requirement. First, he argues that it is needed in order to preserve a transitivity (or 'Xerox') principle, such that if *A* carries the information that *B*, and *B* carries the information that *C*, then *A* carries the information that *C*; he proceeds to argue that preserving this principle is crucial if we want information to 'flow' through a series of structures. Second, he argues that probabilities of 0 or 1 are needed if we want to preserve a conjunction principle, such that if *A* carries the information that *B*, and it also carries the information that *C*, then it carries the information that *B* & *C*. I agree with Dretske that the transitivity and conjunction principles are worth preserving, and that rejecting them in order to make room for error is a high price to pay. The account I develop in subsequent sections is intended to allow for misrepresentation while preserving both principles.

always be distinguished from the real, full-blooded, falsity-allowing ‘non-natural’ meaning of language and thought. But Skyrms openly repudiates this distinction[10], and I think he is right to do so. ‘Non-natural’ meaning must arise somehow or other from natural meaning, so there must be *some* way in which the gap between them can be bridged. And although Skyrms’s account does not take us all the way, it provides a useful platform on which to build.

4. Making room for misrepresentation

4.1 Two strategies

There are two traditional strategies for tackling the problem of error that might conceivably come to the aid of Skyrms’s framework, both of which are inspired by Dennis Stampe’s (1977) seminal discussion of the problems and prospects of causal theories of content. Both involve making the relationship between informational and propositional content a little bit looser; so that, while the propositional content of a signal is still *related* to the information it carries, the relationship is no longer one of immediate deductive entailment.

The first strategy is to augment our basic information-theoretic account with a notion of *function*.[11] We can then say, in broad terms, that the propositional content of a signal depends not on the information it actually carries, but rather on the information it has the function of carrying (see Stampe 1977; Dretske 1981, 1986, 1988; Neander 1995[12]). We might say, for instance, that a particular type of vervet alarm call represents the proposition

---

[10] ‘H. Paul Grice distinguished between *natural* and *non-natural* meaning. Natural meaning depends on associations arising from natural processes. I say that all meaning is natural meaning’ (Skyrms 2010a, p. 1).

[11] The relevant notion of function may be spelled out in developmental (e.g. Dretske 1981) or evolutionary terms (e.g. Neander 1995); the overarching strategy is the same either way.

[12] Of course, the ‘teleosemantics’ programme is also closely associated with Millikan (1984) and Papineau (1984). But Millikan and Papineau see the notion of biological function as providing as an *alternative* to information-theoretic semantics, rather than as providing a means of augmenting the information-theoretic approach to accommodate error. I do not discuss these ‘consumer-based’ variants of teleosemantics in this paper.

that there is a leopard nearby in virtue of the fact that its function is to shift the probabilities of all other possibilities to zero, even if in practice it often fails to do so (owing to the occasional presence of cheetahs, and suchlike). In cases where the alarm call is emitted in the absence of a leopard, the signalling system is malfunctioning[13]: it still represents the proposition that a leopard is nearby, but its propositional content is false.

I will not pursue this route here. One reason for this is that I do not want to go over old ground: the advantages and drawbacks to this kind of response to the problem of error are already well understood, having been explored extensively over the past four decades.[14] But there are more substantive reasons too, for I think the functional strategy encounters particular problems when our aim is to construct a non-derivative account of the content of signals between animals, rather than, say, perceptual representations within an animal. For the guiding intuition behind the functional strategy is that we can account for misrepresentation as a form of *malfunctioning*: that misrepresentation happens when signals fail to indicate what they are supposed to indicate. This seems plausible enough in the context of perceptual representation, since it is plausible to think of errors in perception as a form of malfunctioning. In the context of animal signalling, however, this guiding intuition is harder

---

[13] This would constitute malfunctioning only on the assumption that the alarm call was selected specifically for indicating *leopards*, and not for indicating (say) *large predators*. The functional indeterminacy problem (see footnote 14) threatens to raise its head here, but let's assume for now that the alarm call was specifically selected for indicating leopards. The example is only for illustrative purposes, after all.

[14] See Neander 2012 for a survey. These debates have tended to focus on problems that are shared by all variants of teleosemantics, viz. problems of functional indeterminacy (e.g., Dretske 1986; Sterelny 1990; Griffiths and Goode 1995; Neander 1995) and clashes with intuition in 'swampman'-type scenarios (e.g. Dretske 1995; Neander 1996; Millikan 1996; Papineau 2001). Millikan (1989) poses a further problem that is specific to accounts that combine the notion of function with information theory. Her concern, in a nutshell, is that the proper function of a representation must be an *effect* (specifically, a selected effect) that the representation has on a *consumer* system, and the effect of a representation on a consumer system will be independent of any information it carries about features of the environment. This is a longstanding point of disagreement between Millikan and Neander (see Neander 2013, and Millikan's reply).

to sustain: on the contrary, it seems clear that in this context misrepresentation is *not* always a matter of malfunctioning. Sometimes, a sender can be functioning perfectly well and yet still transmit a false signal as a means of promoting its inclusive fitness interests, often at the expense of the receiver.[15] These cases of false signalling cannot be explained as a form of malfunctioning, but we would still like to account for their possibility.

A second, related problem is that it is hard to see why any signal would ever have the biological function of carrying *propositional* content, as opposed to informational content that is non-propositional, if we follow Dretske and Skyrms in taking propositional content to involve *ruling out* certain states of the world. For in real biological scenarios, a signalling behaviour will almost never completely *rule out* any state of the world (in the sense of shifting its probability all the way to *zero*), and so it would never be selected for doing so: it would instead be selected for shifting the probabilities of states of the world *just enough* to enable an adaptive response in the receiver (cf. Godfrey-Smith 1991, 1992).

The second broad strategy, as I see it, is to appeal in one way or another to *counterfactuals*. This too draws its inspiration from Stampe (1977), and in particular from his discussion of 'fidelity conditions' (p. 50-4). Stampe develops a causal theory of content on which the content of a representation is determined not by the properties of its *actual* cause, but rather by properties of the object that *would* cause its tokening, if certain independently characterized 'fidelity conditions' were to obtain. He notes, however, that:

> Whether my attempt to naturalize contents will work depends upon whether we can gain a sufficiently narrow and empirically well-grounded conception of these so-called "fidelity conditions." This is crucial, because a statement saying what would cause the representation *under relevant conditions of*

---

[15] Skyrms discusses cases along these lines in Chapter 5 of *Signals*, but nowhere does he explain how a signal can carry *false* content, as opposed to content that is true but less than maximally informative. See Godfrey-Smith 2012, 2013 for further discussion of this point.

fidelity is what a specification of content *is*. (Stampe 1977, p. 51)[16]

Although Stampe's theory of content is not explicitly informational in character, this point carries over to the present context. One way to allow for the possibility of error within an information-theoretic approach to content is to say that the propositional content of a signal depends not on the information that it *actually* carries, but rather on the information that it *would* carry under fidelity conditions *C*. We might then say, for instance, that a particular type of vervet alarm call represents the proposition that there is a leopard nearby in virtue of the fact that it *would* shift the probabilities of all other possible states to zero, if only conditions *C* were to obtain. We could then make room for false tokenings—that is, cases in which the alarm call is emitted in the absence of a leopard—provided conditions *C* fail to obtain in these cases.

The challenge we face, if we take this route, is to specify the relevant fidelity conditions in a principled, precise and illuminating way. These conditions must yield determinate descriptions of content in paradigm cases, and it must be clear why *these* conditions—rather than any others—have a privileged role to play in the determination of content (cf. Fodor 1984). Meeting this challenge is not straightforward: almost forty years after Stampe's original discussion, there is no agreement on how, if at all, we can flesh out the details of the counterfactual strategy he suggests.

4.2 Separating equilibria

It is a challenge, however, that Skyrms's new framework may have the resources to meet. For in explicitly connecting the notion of informational content to the formal apparatus of evolutionary game dynamics, Skyrms gives us some powerful new resources on which to draw for the purposes of specifying fidelity conditions. In this section, I want to offer one way

---

[16] Stampe goes on to characterize these fidelity conditions in functional terms. I want to explore a different possibility here.

of pursuing this strategy—one way in which we might weave together information-theoretic and game-theoretic considerations to provide a counterfactual account of propositional content in signalling systems. I should make it clear from the outset that the account is not perfect: it has some important limitations, to which I will turn in Section 6. But I will argue that, despite these limitations, the account makes real progress with the problem of error.

The key game-theoretic notion on which I intend to draw is that of a *separating equilibrium*.[17] A separating equilibrium is an equilibrium of an evolutionary signalling game, in the sense that it is a stationary point in the replicator dynamics. But it is not just any equilibrium: it is a special type of equilibrium at which there is a *one-to-one mapping from states of the world to signals*. In other words, for each of the *n* relevant states of the world, $S_i \in \{S_1, S_2, \ldots, S_n\}$, there is a signal σ_i such that σ_i is sent by a sender if and only if that sender observes S_i. Obviously, this is an ideal scenario for informative signalling. At a separating equilibrium, signalling is maximally informative—as informative as it could possibly be given the setup of the model.

Unsurprisingly, populations often fail to reach separating equilibria, even in very simple signalling games (Skyrms 2010a). Moreover, it is easy to construct signalling games in which a separating equilibrium does not even *exist* in the state-space, owing to the constraints imposed by the structure of the game. So if we refused to ascribe content to signals unless the population in question was *actually at* a separating equilibrium, we would very rarely ascribe any content at all. That, however, it not my proposal. What matters for my purposes is that, even if a population is *not* at a separating equilibrium, it often makes sense to ask which of the various possible separating equilibria the population *would* be most

---

[17] Like many other game-theoretic notions, this notion has been imported to biological game theory from economics, where the one-to-one mapping in question is usually that between the price of a product and its quality (e.g. Rothschild and Stiglitz 1976; Milgrom and Roberts 1982; Wolinsky 1983; Easley and O'Hara 1987; Allen and Faulhaber 1989).

likely to arrive at, if it were to arrive at one. In other words, it often makes sense to ask: 'what is the *nearest* separating equilibrium to the actual state of this population?'

The meaning of 'nearest' here is somewhat context-dependent. In particular, we need to distinguish three sorts of cases. In the first sort, the population actually evolves to a separating equilibrium; in these cases, the 'nearest' separating equilibrium is simply the population's actual state. In the second sort of case, we find that although one or more separating equilibria exist in the model, the population has not evolved to any of them. In these cases, the 'nearest' separating equilibrium is (roughly) the one that is nearest to the population's current state in the state-space of the model.[18] In the third, most troublesome sort of case, no separating equilibrium even exists in the state-space. Sometimes, this will mean that there is no determinate fact of the matter as to the 'nearest' separating equilibrium. But in other cases, we find that, even though no separating equilibrium exists *given the current parameter values*, a small intervention on the model parameters would create one. In these cases, the 'nearest' separating equilibrium is the nearest in the *parameter-space* rather than the state-space: it is that separating equilibrium which would take the smallest intervention on the model parameters to create.

4.3 Propositional content revisited

My proposal is that we employ the notion of the 'nearest separating equilibrium' to characterize the 'conditions *C*' that (if the counterfactual strategy is on the right lines) help determine the propositional content of a signal. What I suggest, in short, is that the

---

[18] This still leaves room for ambiguity. Is the 'nearest' equilibrium the one that would require the smallest change in population frequencies to produce, or the one that would take the least time to arrive at by the shortest route, or the one that has a basin of attraction closest to the population's actual state? There are no doubt further possible measures of dynamical 'distance'. Hopefully, there will be many cases in which all reasonable measures agree on the nearest separating equilibrium, in which case we need not choose between them. When they disagree, we can choose an arbitrary stipulation of the concept of 'nearest' if we want to, but perhaps the more principled response is to say that there is no determinate answer to the question of which separating equilibrium is the 'nearest' in these cases.

propositional content of a signal in a population depends not on the information that it *actually* carries, but rather on the information that it *would* carry *at the separating equilibrium nearest to the population's current state*. I will refer to this proposal as the 'nearest separating equilibrium' (or 'NSE') account of the relationship between informational and propositional content. We can spell it out more precisely as follows:

> **NSE:** Let σ denote a particular type of signal in a signalling game. Let $\mathbf{X}$ denote the actual frequency distribution of signalling strategies in the population at time *t*, and let $\mathbf{X}^*$ denote the counterfactual frequency distribution of signalling strategies at the separating equilibrium nearest to $\mathbf{X}$ in the state-space (or, if appropriate, the parameter-space) of the model. Finally, let $\mathbf{I}^*_\sigma$ denote the informational content that σ *would* carry at $\mathbf{X}^*$. *The propositional content of* σ_i *at* $\mathbf{X}$ *is the disjunction of states compatible with* $\mathbf{I}^*_\sigma$.

Suffice to say, NSE is more baroque than Skyrms's original, seductively simple account of propositional content in signalling games. What do we gain from this additional complexity? What we gain, I think, is an account that yields determinate and reasonable attributions of propositional content across a wide range of cases while allowing for the possibility of misrepresentation.

4.4 The possibility of misrepresentation

How exactly does NSE allow for the possibility of misrepresentation? The short answer is it does so in the much same way as any account that employs some version or other of Stampe's counterfactual strategy (cf. Section 4.1). But perhaps a more detailed explanation will be helpful. Here, I will explain in the abstract how NSE makes room for misrepresentation; in the next section, I will illustrate this with a concrete example.

In Section 3, we saw that Skyrms's account falls foul of the problem of error because it implies that two conditions—Conditions A and B—must be satisfied for a signal to carry false propositional content, and these conditions cannot be satisfied simultaneously. NSE too implies that a signal σ carries false content when two conditions are satisfied. First, it has to be the case that $\mathbf{I}_{\sigma}^{*}$ (the informational content that σ *would* carry at the nearest separating equilibrium) rules out a state of the world, S_x, in the sense that it shifts its objective probability to zero (call this Condition A*). If Condition A* is satisfied, then NSE will ascribe a propositional content that implies ~S_x to the signal. Second, our original Condition B also has to hold: it also has to be the case that, at on at least one actual occasion, σ is sent when S_x is the true state of the world, so that the propositional content of σ is false on at least one occasion. The crucial point is that, at least in principle, Conditions A* and B *can* be satisfied simultaneously. For Condition B requires that $\mathrm{P}\left(\mathrm{S}_{x} \middle| \sigma\right) \neq 0$ in the *actual* population, whereas Condition A* requires $\mathrm{P}\left(\mathrm{S}_{x} \middle| \sigma\right) = 0$ in the population *at the nearest separating equilibrium*. Except in the special case in which the population is actually at a separating equilibrium, these two requirements are compatible. Hence, in the abstract at least, NSE makes room for misrepresentation.

One might still reasonably wonder, however, how this plays out in practice. Does the NSE account make room for misrepresentation in biologically plausible scenarios? And does it underwrite intuitively reasonable ascriptions of propositional content in such cases? I turn to these questions in the next section.

5. An illustration: the Philip Sidney game

To illustrate the NSE account in action, I will apply it in a simple but influential game-theoretic model of animal communication: John Maynard Smith's (1991) 'Philip Sidney' game, developed and extended by Rufus Johnstone and Alan Grafen (1992).[19] The

[19] Philip Sidney was an Elizabethan aristocrat and poet who was mortally wounded at the Battle of Zuphen. The legend goes that, as he lay dying, he handed his water to another wounded soldier, remarking 'thy necessity is yet greater than mine'. Following Johnstone and Grafen (1992), the game is

significance of the game is that, unlike Lewis's original (1969) signalling games, the interests of senders and receivers are not perfectly aligned: there is a *partial conflict of interest* that makes informative signalling difficult to stabilize.

Here is the basic setup.[20] There are two agents: a sender and a donor. The *sender* is in one of two states, healthy (T_1) or needy (T_2), and signals *p*% of the time when healthy and *q*% of the time when needy.[21] Since these percentages can vary continuously, the sender has a continuous range of strategic options in two dimensions. It could, for instance, signal 100% of the time when needy and 50% of the time when healthy; or 50% of the time when needy and 30% of the time when healthy; and so on. The *donor* receives a signal if one is sent, and must choose how to respond. It donates water *x*% of the time when it receives a signal, and donates water *y*% of the time when it does not receive a signal. It too, therefore, has a continuous range of possible strategies in two dimensions. The sender and donor are related to some degree *r*, and this reflects the degree to which their inclusive fitness interests are aligned. Four further features of the setup are important: signalling imposes a fixed cost *c* on the sender, while donating water imposes a fixed cost *d* on the donor. Meanwhile, receiving water confers a benefit *a* on needy beneficiaries and a benefit *b* on healthy beneficiaries, where $a > b$. This *differential* benefit to needy beneficiaries is crucial if informative signalling is to be stabilized.

The equilibria of the Philip Sidney game have been extensively studied (see e.g. Maynard Smith 1991; Johnstone and Grafen 1992; Huttegger and Zollman 2010; Zollman et al. 2013). The overall message from these studies is that changes to the parameter values of

---

sometimes known as the 'Sir Philip Sidney' game, though this was not Maynard Smith's original title. The original title is more apt, since a knighthood is a living honour which lapses upon death.

[20] For the particular version of the game studied in Huttegger and Zollman 2010.

[21] In this scenario, unlike in the case of alarm calls, the sender is not necessarily 'observing' a state of the world, since it may have non-observational sensitivity to its own health. But we can say that the sender 'observes' its own health if we want to, since the mechanism by which a sender comes to detect the relevant state of the world makes no difference to the formalism.

the model significantly affect the number and nature of equilibria that exist. The equilibria vary depending on the (i) the cost of signalling, *c*; (ii) the cost of donating, *d*; (ii) the differential benefit to needy individuals, *a* – *b*; and (iii) the extent to which the agents' inclusive fitness interests are aligned, *r*. Here are some notable equilibria, all of which exist somewhere in the parameter-space:

> *The classic separating equilibrium:* The sender signals if and only if it is needy; the donor donates if and only if it receives a signal.
>
> *An unorthodox separating equilibrium:* The sender signals if and only if it is *healthy*; the donor donates if and only if it does *not* receive a signal.
>
> *A hybrid equilibrium:* The sender signals 100% of the time if needy. If healthy, the sender mixes between signalling (with probability α) and not signalling (with probability $1 - \alpha$). The donor never donates if it does not receive a signal. If it does receive a signal, it mixes between donating (with probability β) and not donating (with probability $1 - \beta$).

The existence of hybrid equilibria for certain combinations of parameter values in the Philip Sidney game is proven by Simon Huttegger and Kevin Zollman (2010). Moreover, they prove that hybrid equilibria and separating equilibria never co-exist in the dynamics: changing the parameter values so as to create a separating equilibrium destroys any hybrid equilibrium, and vice versa. This is because a separating equilibrium exists only if $a \geq c + rd \geq b$ (for the classic separating equilibrium) or $a \geq rd - c \geq b$ (for the unorthodox separating equilibrium), whereas a hybrid equilibrium exists only if $b - rd > c$. No single specification of parameter values can possibly satisfy both conditions.

We need not concern ourselves with any further technical details, because the question I want to consider here is an interpretative one: *at the hybrid equilibrium, what do* *the senders' signals mean?* For, as Zollman et al. (2013) note, it is intuitive to think that, when the population is at a hybrid equilibrium, the needy individuals are *truthfully* signalling their need, whereas a fraction α of healthy individuals are *falsely* signalling that they too are needy: 'in plain English, this means that the [healthy] sender sometimes "lies" and is honest at other times' (Zollman et al. 2013, p. 4).

The scare quotes around 'lies' reflect the fact that it is difficult to find any basis for this intuition within an information-theoretic framework. Real lying, after all, requires false propositional content, and information-theoretic accounts of content (such as that of Skyrms) tend to leave no room for this. The most Skyrms's account could say in this case is that, in signalling to a donor, a healthy sender upshifts the probability that it is needy. The sender can therefore (in this minimal technical sense) be said to be signalling 'deceptively' (Skyrms 2010a, Ch. 5). But Skyrms's account does not support any ascription of false propositional content to the healthy sender's signal. Indeed, Skyrms's account will not support any ascriptions of propositional content *at all* in this case, since the senders' signals do not downshift any probabilities to zero.

The NSE account, however, does support ascriptions of propositional content to the signals at a hybrid equilibrium. Moreover, it vindicates our intuitive ascriptions of content. This is because, for any hybrid equilibrium in the Philip Sidney game, there is a determinate fact of the matter about the nearest separating equilibrium in the parameter space—the separating equilibrium it would take the smallest intervention on the parameter values to create. Assuming we start with parameter values that are biologically plausible, the nearest separating equilibrium will usually be the classic separating equilibrium, because the unorthodox separating equilibrium only exists at parameter values that are rather biologically implausible (r must be very high, and/or d must be much greater than b). Hence, the nearest separating equilibrium will usually be one at which senders would signal only if needy; it would therefore be one at which the signal would upshift the probability that the sender is needy to 1, and would downshift the probability that the sender is healthy to 0. The NSE account thus supports the intuition that, at the hybrid equilibrium, the signal represents the proposition that the sender is needy, and that the signal carries *false* propositional content when the sender is healthy.

6. The partition problem

Here, then, is the good news: the NSE account of the relationship between propositional and informational content in signalling systems makes room for the possibility of misrepresentation, and delivers intuitively reasonable ascriptions of false propositional content to signals at the hybrid equilibria of the Philip Sidney game. I conceded from the outset, however, that the account it not perfect, and it is now time to turn, briefly, to its limitations.

The Philip Sidney game is a useful illustrative case, because it allows us to see how questions like 'what *would* this signal have indicated at the nearest separating equilibrium?' can sometimes be given determinate answers. But one might worry how often the answers are going to be determinate. Sometimes we will have whole classes of separating equilibria, and no easy way to tell which is 'nearest' to the population's current state. In other cases, there will be no separating equilibrium and no easy way of creating one by intervening on the parameter values. The outcome will be the same: various separating equilibria might be conceivable given extreme enough interventions, but none will be unambiguously 'nearest' to the population's current state.

What should we say about such cases? One option would be to insist that the NSE account will tell us what the propositional content of a signal is whenever there is any determinate content to be had: conversely, if there is no determinate fact of the matter as to the nearest signalling equilibrium, then there is no determinate propositional content either. This, however, seems a little extreme. Sometimes, there is no unambiguously 'nearest' separating equilibrium, and yet no intuitive reason for withholding ascriptions of content.

Suppose, for instance, we have a model with 100 relevant states of the world and 99 signals; and suppose that 98 of the signals reliably indicate a particular state of the world, while the remaining signal indicates that one of the remaining two states obtains. Strictly speaking, this would not be a 'separating equilibrium', because there would be no *one-to-one* mapping from signals to states. Nor would there be an unambiguous 'nearest' separating equilibrium in the parameter-space: we could bring many separating equilibria into existence by adding an extra signal to the model, but this would not determine *which* separating equilibrium the population would be most likely to arrive at (e.g. would the new signal come to indicate state 99 or state 100?). And yet, in a case in which we have a *near*-perfect mapping from states to signals, it seems overly restrictive to conclude that there is no propositional content merely because there is no determinately nearest separating equilibrium. The NSE account seems to struggle with these cases.

One response to this sort of example would be to maintain that we *can* still have a separating equilibrium when we appear to have more states than signals, if we are prepared to individuate states of the world at a coarser grain of analysis. For instance, our model setup may suppose that State 99 and State 100 are relevantly distinct, but if no signal discriminates between them, then perhaps we should rethink: perhaps we should reclassify them as a single state.

This response, however, merely leads us directly to a broader, deeper problem: how are states of the world to be individuated in any principled way, outside the context of simple formal models? The notion of a 'separating equilibrium', after all, only makes sense *relative to a specification of a set of relevant states of the world*. In evolutionary game theory, it is usually taken for granted that we can help ourselves to such a specification—that specifying the set of relevant states is part of the modeller's tacit skill. But it is reasonable to ask for further justification if we are to base a theory of *content* on these foundations. We need a partition of states of the world that is not just a *reasonable* one for most purposes, but the *right* one for the specific purpose of individuating semantic contents.

Suppose, for instance, we are modelling alarm calls in vervet monkeys. Do we model the presence of a leopard and the presence of a cheetah as discrete states of the world? If so, then the actual population is not at a separating equilibrium, because some alarm calls are tokened in the presence of both leopards and cheetahs. But should we instead model the presence of either predator as a single state—the presence of a *large feline predator*, perhaps? At this more coarse-grained partitioning of states of the world, the actual population might well be at a separating equilibrium. Hence, the notion of a separating equilibrium is inherently partition-relative. The resultant concern that is that we could re-describe virtually *any* signalling equilibrium of any population as a separating equilibrium if we were prepared to partition states of the world at a sufficiently coarse grain of analysis. Likewise, we could make a separating equilibrium virtually impossible for any population to attain by partitioning states of the world at an excessively fine grain.

We can call this the 'partition problem' for game-theoretic approaches to content.[22] I have no solution to offer here. Note, however, that the problem is not *specific* to the NSE account, or even to accounts that employ the apparatus of game theory. It is a variant of a general problem that afflicts *any* information-theoretic approach to content. Formalizations of information theory almost invariably require a set of relevant possibilities to be specified by the theorist, and no information-theoretic account of content will ever yield plausible ascriptions of content unless these possibilities are individuated at the right grain of analysis (cf. Dretske 1981, 1999; Cohen 2004). Solving this problem remains an important task for future work.[23]

---

[22] The partition problem is clearly reminiscent of the 'disjunction problem', and might be regarded as the way in which one aspect of that problem resurfaces for the NSE account. But note that the problem here is purely one of finding a principled partition of states of the world. Relative to any given partition, the NSE account is able to distinguish between true and false tokenings of a signal; so this aspect of the traditional 'disjunction problem' does not resurface.

[23] One option here is to individuate states by their payoffs: if two putative states yield identical payoffs under all circumstances, then they are identical for the purposes of content ascription. One drawback to

7. Outlook

The NSE account of propositional content has notable successes. It seems unlikely, however, that it tells us the whole story of how propositional content arises from the flow of information. Undoubtedly, more work remains to be done. Yet I hope the account's successes are enough to show that the programme it exemplifies is viable. A naturalistic approach that interweaves elements from information theory and game theory is surely no dead end, even when our explanatory target is the origin of Meaning with a capital M: the representation (and misrepresentation) of propositions. For when their strengths are combined, these two theoretical frameworks provide us with a formidable set of resources. By employing them in an integrated fashion, we can start to see how misrepresentation—the mark of 'non-natural' meaning—could arise from natural, informational connections between signals and states of the world.

Acknowledgements

Versions of this article were presented between March and June 2013 at the University of Cambridge, the University of Bristol and Duke University. I thank the audiences at these events for their helpful questions and comments, and I thank Peter Godfrey-Smith, Françoise Longy, Manolo Martínez, Nick Shea, Ulrich Stegmann and an anonymous referee for their detailed and insightful comments on the manuscript.

References

Adams F, Aizawa K (2010) Causal theories of mental content. In: Zalta E (ed.) *The Stanford encyclopedia of philosophy* (Spring 2010 edition). URL: http://plato.stanford.edu/archives/spr2010/entries/content-causal/ (Last accessed: 26/01/14)

---

this move is that, in real ecological scenarios, payoffs are likely to depend in an extremely fine-grained way on environmental differences, suggesting that few natural populations would even approximate a separating equilibrium by this criterion. But this remains an avenue worthy of further exploration.

Adams F, Beighley SM (2013) Information, meaning and animal communication. In: Stegmann UE (ed.) *Animal communication theory: information and influence*. Cambridge University Press, Cambridge, pp. 399-420.

Allen F, Faulhaber GR (1989) Signalling by underpricing in the IPO market. *Journal of Financial Economics* 23:303-323.

Cheney DL, Seyfarth RM (1990) *How monkeys see the world: inside the mind of another species.* University of Chicago Press, Chicago.

Cohen J (2004) Information and content. In: Floridi L (ed.) *The Blackwell guide to the philosophy of computing and information*. Blackwell, Oxford, pp. 215-227.

Crane T (2003) *The mechanical mind: a philosophical introduction to minds, machines and mental representation* (2nd edition). Routledge, London.

Di Bitetti MS (2003) Food-associated calls of tufted capuchin monkeys (Cebus appella nigritus) are functionally referential signals. *Behaviour* 140:565-592.

Dretske F (1981) *Knowledge and the flow of information*. MIT Press, Cambridge MA.

Dretske F (1986) Misrepresentation. In: Bogdan R (ed.) Belief: form, content, and function. Oxford University Press, New York pp. 17-36.

Dretske F (1988) *Explaining behavior: reasons in a world of causes*. MIT Press, Cambridge MA.

Dretske F (1995) *Naturalizing the mind*. MIT Press, Cambridge MA.

Dretske F (1999) Informational semantics. In: Wilson RA, Keil FC (eds.) *The MIT encyclopedia of the cognitive sciences*. MIT Press, Cambridge MA, pp. 406-408.

Easley D, O'Hara M (1987) Price, trade size and information in securities markets. *Journal of Financial Ecomomics* 19:69-90

Evans CS (1997) Referential signals. In: Owings DH, Beecher MD, Thompson NS (eds) *Perspectives in ethology*. Plenum, New York, pp. 99-143.

Evans CS, Evans L (2007) Representational signalling in birds. *Biology Letters* 3:8-11.

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

Fodor JA (1987) *Psychosemantics: the problem of meaning in the philosophy of mind*. MIT Press, Cambridge MA.

Fodor JA (1990) Information and representation. In: Hanson P (ed.) *Information, language, and cognition*. University of British Columbia Press, Vancouver, pp. 175-190.

Gill SA, Bierema AMK (2013) On the meaning of alarm calls: a review of functional reference in avian alarm calling. *Ethology* 119:449-461.

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

Godfrey-Smith P (1991) Signal, decision, action. *Journal of Philosophy* 88:709-722.

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

Godfrey-Smith P (2012) Review of Skyrms (2010a). *Mind* 120:1288-1297.

Godfrey-Smith P (2013) Signals, icons, and beliefs. In: Ryder D, Kingsbury J, Williford K (eds), *Millikan and her critics*. Wiley-Blackwell, Chichester, 41-62.

Grice HP (1957) Meaning. *Philosophical Review* 66:377-388.

Griffiths P, Goode PE (1995) The misuse of Sober's selection for/selection of distinction. *Biology and Philosophy* 10:99-107.

Huttegger SM, Zollman KJS (2010) Dynamic stability and basins of attraction in the Sir Philip Sidney game. *Proceedings of the Royal Society of London Series B: Biological Sciences* 277:1915-1922.

Johnstone RA, Grafen A (1992) The continuous Sir Philip Sidney game: a simple model of biological signalling. *Journal of Theoretical Biology* 156:215-234.

Lewis DK (1969) *Convention: a philosophical study*. Harvard University Press, Cambridge MA.

Macedonia JM (1990) What is communicated in the antipredator calls of lemurs: evidence from antipredator call playbacks to ringtailed and ruffled lemurs. *Ethology* 86:177-190.

Macedonia JM, Evans CS (1993) Variation among mammalian alarm call systems and the problem of meaning in animal signals. *Ethology* 93:177-197.

Manser MB (2001) The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proceedings of the Royal Society of London Series B: Biological Sciences* 268:2315-2324.

Manser MB (2009) What do functionally referential alarm calls refer to? In: Dukas R, Ratcliffe J (eds.) *Cognitive ethology II*. University of Chicago Press, Chicago, pp. 229-248.

Manser MB, Bell MB, Fletcher LB (2001) The information that receivers extract from alarm calls in suricates. *Proceedings of the Royal Society of London Series B: Biological Sciences* 268:2485-2491.

Maynard Smith J (1991) Honest signalling: the Philip Sidney game. *Animal Behaviour* 42:1034-1035.

Milgrom P, Roberts J (1982) Limit pricing and entry under incomplete information: an equilibrium analysis. *Econometrica* 50:443-459.

Millikan RG (1984) *Language, thought and other biological categories*. MIT Press, Cambridge MA.

Millikan RG (1989) Biosemantics. *The Journal of Philosophy* 86:281-297.

Millikan RG (1996) On swampkinds. *Mind and Language* 11:103-117.

Neander K (1995) Misrepresenting and malfunctioning. *Philosophical Studies* 79:109-141.

Neander K (1996) Swampman meets swampcow. *Mind and Language* 11:118-129.

Neander K (2012) Teleological theories of mental content. In: Zalta E (ed.) *The Stanford encyclopedia of philosophy* (Spring 2012 edition). URL: http://plato.stanford.edu/archives/spr2012/entries/content-teleological/ (Last accessed: 26/01/14)

Neander K (2013) Towards an informational teleosemantics. In: Ryder D, Kingsbury J, Williford K (eds), *Millikan and her critics*. Wiley-Blackwell, Chichester, 21-40

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

Papineau D (1984) Representation and explanation. *Philosophy of Science* 51:550-572

Papineau D (2001) The status of teleosemantics, or how to stop worrying about swampman. *Australasian Journal of Philosophy* 79:279-289.

Rendall D, Owren MJ, Ryan MJ (2009) What do animal signals mean? *Animal Behaviour* 78:233-240.

Rendall D, Owren MJ (2013) Communication without meaning or information: abandoning language-based and informational constructs in animal communication theory. In: Stegmann UE (ed.) *Animal communication theory*. Cambridge University Press, Cambridge, pp. 151-188.

Rothschild M, Stiglitz J (1976) Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. *The Quarterly Journal of Economics* 90:629-649.

Seyfarth RM, Cheney DL, Marler P (1980) Vervet monkey alarm calls: semantic communication in a free-ranging primate. *Animal Behaviour* 28:1070-1094.

Shannon CE (1948) A mathematical theory of communication. *Bell System Technical Journal* 27:379-423.

Shannon CE, Weaver W (1949) *The mathematical theory of communication*. University of Illinois Press, Chicago.

Slocombe KE, Zuberbühler K (2005) Functionally referential communication in a chimpanzee. *Current Biology* 15:1779-1784.

Skyrms B (2008) Presidential address: signals. *Philosophy of Science* 75:489-500.

Skyrms B (2009) Evolution of signalling systems with multiple senders and receivers. *Proceedings of the Royal Society of London Series B: Biological Sciences* 1518:771-779.

Skyrms B (2010a) *Signals: evolution, learning, and information*. Oxford University Press, Oxford

Skyrms B (2010b) The flow of information in signalling games. *Philosophical Studies* 147:155-165.

Skyrms B (2012) Learning to signal with probe and adjust. *Episteme* 9:139-150.

Skyrms B, Huttegger SM (2013) Emergence of a signaling network with probe and adjust. In: Sterelny K, Joyve R, Calcott B, Fraser B (eds.) *Cooperation and its evolution*. MIT Press, Cambridge MA, pp. 265-274.

Stampe DW (1977) Toward a causal theory of linguistic representation. *Midwest Studies in Philosophy* 2:42-63.

Sterelny K (1990) *The representational theory of mind: an introduction*. Blackwell, Oxford.

Townsend SW, Manser MB (2013) Functionally referential communication in mammals: the past, present and future. *Ethology* 119:1-11.

Zollman KJS, Bergstrom CT, Huttegger SM (2013) Between cheap and costly signals: the evolution of partially honest communication. *Proceedings of the Royal Society of London Series B: Biological Sciences* 280:20121878.

Wolinsky A (1983) Prices as signals of product quality. *Review of Economic Studies* 50:647-658.

Zuberbühler K (2000) Referential labelling in wild Diana monkeys. *Animal Behaviour* 59:917-927.