ORIGINAL ARTICLE

Attribution of Information in Animal Interaction

Stephen Francis Mann[1]

Abstract

This article establishes grounds on which attributions of information and encoding in animal signals are warranted. As common interest increases between evolutionary agents, the theoretical approach best suited to describing their interaction shifts from evolutionary game theory to communication theory, which warrants informational language. The take-home positive message is that in cooperative settings, signals can appropriately be described as transmitting encoded information, regardless of the cognitive powers of signalers. The canonical example is the honeybee waggle dance, which is discussed extensively in the second and third sections. The take-home negative message is that signals are not always a consequence of coadaptation. The communication theory approach is just one end of a continuum explored more thoroughly by evolutionary game theory. The fourth and fifth sections explore this wider framework, as well as overturning some widely held misconceptions about information theory.

Keywords Animal communication · Behavioral ecology · Communication theory · Evolutionary game theory · Teleosemantics

Introduction: The Information Debate in Behavioral Ecology

This article establishes grounds on which attributions of information in animal signals are warranted. Informational language is most appropriate when signaling behavior is the result of coadaptation. Animal interactions are more generally understood through the lens of evolutionary game theory. As common interest increases between evolutionary agents, the theoretical approach best suited to describing their interaction shifts from game theory to communication theory, which warrants informational language.

The article proceeds in a slightly unusual form. The remainder of this section is given over to outlining the context of the controversy surrounding information concepts in behavioral ecology. Then, a very early information-theoretic approach to animal signaling, which provides an informational measurement of the honeybee waggle dance, is described in the next section. In the following section some recent mathematical objections to that early work are

rebutted. In the next section on "Wider Considerations," broader skeptical arguments concerning the use of informational concepts in animal communication theory are considered. Although the outlook is generally positive, skeptical arguments are only partly rejected. Finally, in the last section before the conclusion, prospects for extending the use of informational concepts to other animal signals are considered. Although several lines of skepticism have been conflated and some are inappropriate, there remain grounds for caution.

Overall, several lines of argument are presented in favor of the use of information theory for studying certain biological communication systems. Throughout, a positive proposal is advanced for the interpretation of information in biological signals. Information is a measure of the accuracy with which a receiver's goal is achieved, or equivalently, the extent to which a receiver's biological capacity is realized. In our case study, which focuses on the waggle dance of the honeybee *Apis mellifera*, the goal is finding food. Insofar as receiver bees have a greater probability of finding food after having followed a dance, information is transmitted by dancing bees. It is this notion of transmission that constitutes grounds for application of the information concept, and related mathematical tools, in studies of biological signaling.

✉ Stephen Francis Mann
  stephenfmann@gmail.com

1 School of Philosophy, Australian National University, Canberra, ACT, Australia

The take-home positive message is that the communication theory approach is a special case that emerges as signaler and receiver interests become aligned. These considerations give grounds for optimism about the use of informational concepts and measurements in behavioral ecology and, more generally, biology. Paradigmatic cases of information transmission are those in which the form of the signal is designed by coadaptation of signaler and receiver. In these cases, signal structure can appropriately be described in terms of a "code." Consequently, the take-home negative message is that signals are not always shaped exclusively by coadaptation. The communication theory approach is just one end of a continuum explored more thoroughly by evolutionary game theory. As a result, signal form may not always be explicable by reference to encoding, at least not as it is now understood. Whether or not signals are manipulative or coadaptive will be an empirical question in each case.

Background

By the late 1970s, animal signals were typically defined in terms of information. Signal evolution was thought to mainly proceed via ritualization of cues. Since ritualization is a process of coadaptation, signals were viewed as primarily cooperative, despite the possibility of deception and exploitation. In the face of this orthodoxy, Dawkins and Krebs (1978) offered a different approach. They proposed that communication be seen as one animal controlling another, typically by exploiting perceptual mechanisms designed for other purposes. Due to the individualistic operation of natural selection, cooperation and coadaptation are rare (Dawkins and Krebs 1978, p. 289). Consequently, we should adopt an individualistic account of communication, one that places the signaler front and center.

More recently, in a series of individual and joint papers, Michael J. Owren, Drew Rendall, and Michael J. Ryan (hereafter ORR) present significant challenges to the use of informational concepts in animal communication studies (Rendall et al. 2009; Owren et al. 2010; Rendall and Owren 2013; Ryan 2013). Information, they claim, is an insubstantial metaphor that cannot do explanatory work and often misleads us as to the nature of signaling behavior. Several behavioral ecology texts go so far as to define communication in terms of information, without ever fully explicating the latter (see Rendall et al. (2009, Table 1) for examples). Instead, ORR propose a definition of signaling akin to that of Dawkins and Krebs, in terms of the influence one animal exerts over another.[1] Further skepticism is advanced by

Sarkar (2013), who discusses the honeybee waggle dance along with wider issues of animal signaling. Sarkar claims different informational concepts have been conflated in the literature, and that as a result informational measurements do not capture the quantities claimed by those who employ them.

It is in the light of this skepticism that the present article is offered. The honeybee waggle dance is one of the most thoroughly studied animal communication systems in nature. It is characterized by an unusually high level of common interest which, I argue below, renders it suitable for informational analysis. I shall first describe, in the next section, one informational measurement of the honeybee waggle dance and respond in the following section to preliminary mathematical objections as to its validity. I will then face the problems raised by ORR, Sarkar, and others that apply more forcefully when we consider instances of communication not characterized by high common interest.

The Honeybee Waggle Dance

"The" honeybee waggle dance is rather a family of communicative behaviors performed by all seven species of honeybee (Beekman et al. 2015, p. 1). Six species are native to Southeast Asia, but the seventh—the Western honeybee, *A. mellifera*—is widespread in Europe, Africa, and Asia and has been domesticated since antiquity, making it uniquely amenable to scientific study. Western honeybees perform the waggle dance when foraging and searching for new nest sites. The works discussed here focus on *A. mellifera* foraging at artificial feeders and performing the waggle dance on a vertical surface inside the hive (see, for example, von Frisch (1950, pp. 76–77)). It is in this context that I use the term "honeybee waggle dance."

When individual bees discover a valuable food source, it is often beneficial to recruit nestmates as soon as possible. Competition from other foragers, and the brief duration of flowering, set important time constraints. A bee that finds a good flower patch will often try to direct available workers to that location. Indicating direction and distance on the vertical inner surface of a pitch-black hive is no mean feat. Bees nonetheless succeed by performing repeated patterns of figure-eight movement whose detectable features—orientation and duration—correspond to the direction and distance of the food source (see Fig. 1). This is the waggle dance, so called because the bee's body vibrates during the straight portion of the figure-eight run with a frequency

---

[1] Stegmann's edited volume (Stegmann 2013a) characterizes the debate as between information-based and influence-based definitions. However, as pointed out by several entries in that volume, information and influence do not form a strict dichotomy. Here I am primarily

Footnote 1 (continued)

concerned with establishing grounds for the attribution of information, so I make little mention of influence-based definitions.
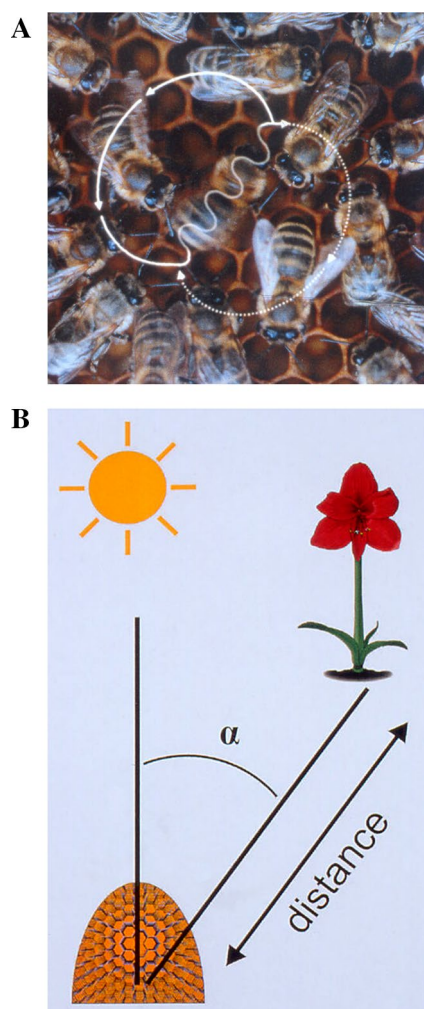
Fig. 1 How the waggle dance indicates the direction of food. Caption from the original image: "A waggle run oriented 45° to the right of 'up' on the vertical comb (**A**) indicates a food source 45° to the right of the direction of the sun outside the hive (**B**). The abdomen of the dancer appears blurred because of the rapid motion from side to side. (Figure design: J. Tautz and M. Kleinhenz, Beegroup Würzburg.)" (Chittka 2004, p. 898)

corresponding to the quality of food discovered. Interested recruits physically follow the dancer to familiarize themselves with the direction and distance indicated, before flying off to locate the relevant food patch.

Informally, the dance is a signal about the location of food. Behavioral ecologists advert to the correspondence between dance and food when explaining foraging behavior of hive recruits (von Frisch 1950, p. 78; Biesmeijer and Seeley 2005, p. 133; Riley et al. 2005, p. 205; Beekman et al. 2015, p. 1). Indeed, much work on honeybees after von Frisch was geared towards demonstrating that these explanations were correct (Gould 1975). That recruits are capable of finding food at a rate better than chance having followed the

relevant dance is an established fact.[2] The present work uses the foraging waggle dance as a case study in the application of information theory to animal communication. We shall assume that recruits who follow the dance find the indicated food source at better than chance and that the dance evolved for this reason.

The next subsection surveys the earliest attempt to measure information transmission in the waggle dance. I then describe and respond to recent objections.

Haldane and Spurway (1954)

In this section I build a positive case for the relevance of informational measurements to selective explanation. The claim is illustrated with a case study of the honeybee waggle dance, whose informational properties were first measured by Haldane and Spurway (1954). I will now describe the aims, methods, and conclusions of that study.

Haldane and Spurway wanted to demonstrate the possibility of deriving informational properties from statistical data. They used honeybee communication as an example because von Frisch had already published the relevant statistics (von Frisch 1948, 1950, 1952). By 1954, mathematical communication theory (MCT) offered a new perspective on this data from an engineering viewpoint. At the heart of the motivation for the study was the possibility of comparison with other animals. It was hoped that the way information is measured, and the generality of the units of information, might allow for the magnitude of information transmission to be compared between communication systems. Indeed, soon after the study was published Wilson (1962) obtained similar statistical and informational data for fire ants, explicitly comparing his results with those of the earlier work. Whether or not this comparison was valid is discussed below in the section "Information, Costs and Benefits."

Although contemporary work on biological information drawing on MCT typically begins by citing Shannon (1948a, b) or Shannon and Weaver (1949), Haldane and Spurway took their leave from Norbert Wiener's *Cybernetics* (Wiener 1948). Wiener's approach emphasizes intra-system control over inter-system communication. Since a honeybee colony may fruitfully be construed as a more or less unified entity with unified goals, it is appropriate to take a cybernetical approach to its behavior.[3]

---

[2] It is less clear how useful the dance is across different contexts, and the purpose for which it originally evolved. Several lines of evidence count in favor of nest site selection as the significant factor (Beekman et al. 2008; I'Anson Price and Grüter 2015). For simplicity I regretfully ignore this possibility.

[3] Haldane and Wiener knew each other personally. For some interesting remarks on their relationship, see Dronamraju's recent biography of Haldane (Dronamraju 2017, pp. 259–260).
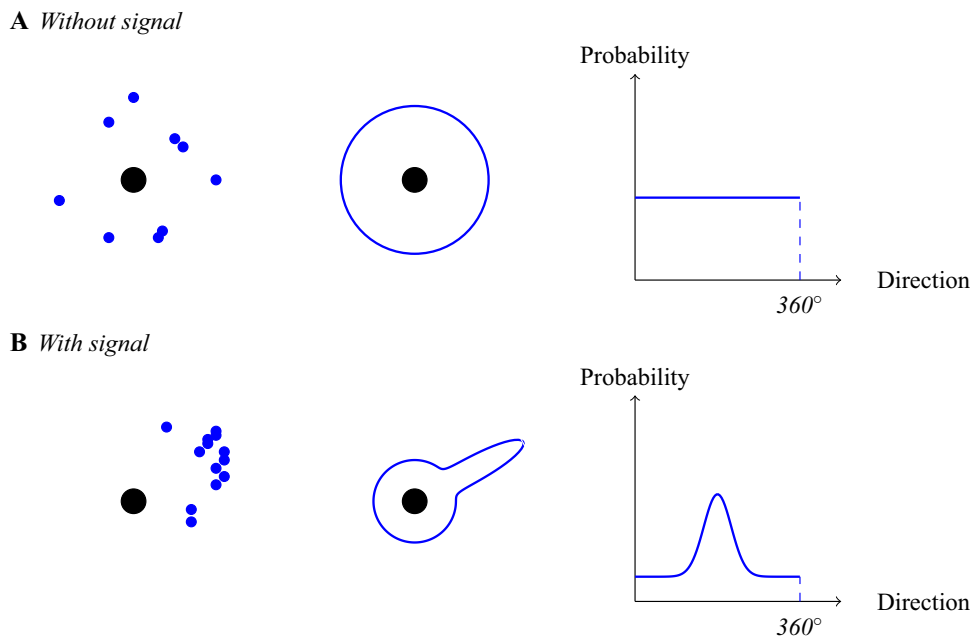
A *Without signal*

Probability

Direction

360°

B *With signal*

Probability

Direction

360°

Fig. 2 How directional information in the waggle dance is measured. **A** Without experience, bees are expected to disperse randomly (small dots; blue in online version) about the hive (black disc). This is represented by a uniform distribution (circle; blue in online version) around the hive. The same distribution in Cartesian coordinates is also presented. **B** After receiving a signal, receiver bees are much more biased toward the direction indicated by the dance. This situation is represented as a Gaussian distribution. Information transmission is measured by subtracting the entropy of **B** from the entropy of **A** (see Eq. 1 for details). Although individual entropies are quantified relative to the coordinate system (here 360 degrees per circle), information rate is not relativized in this way (Shannon 1948b, p. 631).

Haldane and Spurway (1954, p. 255) use just one equation to measure information. It is a formula that converts two pieces of statistical data (representing receiver behavior) into a term for information transmitted by senders. The conversion works as follows. Suppose a group of receivers are about to embark on a foraging run. In the absence of experience they will tend to spread themselves evenly around the foraging domain. Their spatial distribution is represented by a uniform probability distribution (Fig. 2A). This is the first piece of statistical data, generated by the modeling assumption that bees without information forage randomly. Note that this is a continuous distribution. The foraging domain has *not* been segmented into discrete "cells," each with some nonzero probability of being arrived at. Rather, *any* given region—of any size—has a nonzero probability of being visited. Now consider how receiver bees are distributed when they react to a dance. Instead of dispersing randomly throughout the domain, they all fly in roughly the same direction (Fig. 2B). This is the second piece of statistical data, generated by the observed distribution of recruits which is assumed to be Gaussian. This too is a continuous function. The "most popular" direction is the direction of food, but some receiver bees will be a little inaccurate. The conversion equation translates these two distributions into a quantity of information, which can be interpreted as the information provided by senders about the direction of food. It is calculated as follows.

The conversion equation is derived from two formulas taken from Shannon (1948b). These formulas give the entropy of a Gaussian distribution (Shannon 1948b, p. 630) and the information rate of a continuous channel (Shannon 1948b, p. 637). Haldane and Spurway combine the two into a single equation giving *the information rate of a continuous channel when the source entropy is uniform and the conditional entropy is Gaussian*. The most perspicuous form of this equation is due to Wilson (1962, Appendix, p. 156) and I present it here (the units are bits per signal):

$$\text{Information rate} = \log_2 \frac{360°}{\sigma°} - \log_2 \sqrt{2\pi e} \qquad (1)$$

Here, σ represents the standard deviation of the Gaussian distribution. Informally, it measures the "spread" of receiver bees about the food source. Granted that receivers can be represented by a Gaussian distribution, σ is the only statistic required to calculate information rate.[4]

---

[4] The appropriate measure for the Gaussian distribution is presented in Shannon and Weaver (1949, p. 89). Haldane and Spurway cite Wiener (1948, p. 62) who provides a general formulation of information in a continuous distribution and does not appear to discuss the Gaussian case explicitly. For a derivation (and explanation) of the equation used by Haldane and Spurway (1954, p. 255) see Wilson (1962, Appendix).

An equivalent way of conceiving the situation is to take the uniform distribution as a "prior" and the Gaussian as a "posterior." Before the dance, receiver bees are maximally uncertain about the location of food, hence the prior distribution is uniform. After the dance, the receiver bees' uncertainty has decreased. The magnitude of the decrease in uncertainty is a relation between the Gaussian and uniform distributions.[5]

By the method just described, Haldane and Spurway derive a quantity of information about the direction in which food can be found. Substituting the standard deviation of 14.7° (1954, p. 251) into Eq. (1), they conclude that each recruit receives on average 2.5 bits (1954, p. 278) of information about the direction of food. It is worth pointing out that this analysis holds regardless of facts about the cognitive powers of individual bees. On the cybernetic interpretation, information is a measure of how behavioral precision contributes to system goals. Cognitive sophistication can be included or omitted from models of communication, without bearing on the presence of information in signaling systems. The same is true of sender-receiver models in general (Skyrms 2010, p. 44).

Why is this method justified in the case of honeybees? If a colony is an entity with unified evolutionary goals, it is plausible to think senders obtain a payoff for receiver success. Since receiver success increases when information is transmitted, the dance evolves because of the information it provides. In other words, when "information" is construed in terms of accuracy, information rate is selected for. Nature plausibly selects for accuracy under the circumstances bee colonies have historically found themselves in. On this interpretation, selection has led to increased information transmission. This is one route to explaining the increased articulation in the waggle dance over phylogenetically earlier kinds of dance. If articulation supports greater information transmission, and information rate is selected for, then articulation can be selected for its informational consequences. Of course, adaptation is always subject to constraints, and the quantity of information transmitted cannot increase indefinitely (Preece and Beekman 2014). Nonetheless, it is justifiably claimed that the benefits of foraging accuracy historically underpinned selection for greater information transmission in the honeybee waggle dance.

In sum, Haldane and Spurway established that informational measures can be derived from statistical data. They justified their use of a measure taken from MCT by adopting a cybernetic perspective on honeybee colonies. To the extent that a colony's communicative goals are unified, this

approach is justified. Since 1954 more work in a similar vein has appeared (Riley et al. 2005; Beekman et al. 2015; Schürch and Ratnieks 2015). Regardless of differing results, what matters here is whether Haldane and Spurway's model was a good one and whether their interpretation of information was valid. I have so far argued that the answer to both of these questions is yes. The next section responds to two objections.

Initial Objections to Haldane and Spurway

Skepticism of the validity of these results continues (Pfeifer 2006; Sarkar 2013). Even optimistic scholars, as well as those on the fence, cite Pfeifer (2006) as having presented technical challenges to the use of information theory in biology (Reading 2011, p. 149, fn. 6; Stegmann 2013b, p. 143; Wiley 2013, p. 118). I aim to show that although there surely will be technical and methodological hurdles, they are not the ones presented by these objections.

Objection 1: Arbitrary Coordinate Systems

In this section I show how the bee dance is both iconic and continuous and how signals of this kind contain measurable information. This undercuts one skeptical argument against the use of information theory for animal signals in general, and the bee dance in particular, namely that the models we employ to quantify signal information are somehow arbitrary. Discrete models appear to be arbitrary, but continuous models do not suffer from the same defect.

The skeptical argument due to Pfeifer (2006) and repeated by Sarkar (2013, §7.3) runs as follows. Honeybee signals indicate, among other things, the direction of a food source relative to the hive. The quantity of information about direction in a signal depends on how precise the signal is. When modeling bee signals, therefore, the amount of information we will obtain depends on how finely we carve up "direction space." But we have no a priori guidance as to how finely to carve up the space. Different divisions will give different measures of information, but none will have priority. Sarkar suggests we could test for individual bees' perceptual acuity, and divide the space into segments of a size roughly discriminable by the bees themselves. He claims Haldane and Spurway did not do this, and that even if they had, the placement of these individual direction-segments would have been arbitrary. By shifting them a half-step clockwise, we obtain a totally new division of direction space, which will transform our statistical data into different informational measurements. Since no carving of the space is privileged, there can be no privileged conversion of statistical data into informational quantities. As a result, information

---

[5] In fact the von Mises distribution would have been more appropriate (Schürch and Ratnieks 2015). The Gaussian is an acceptable approximation.
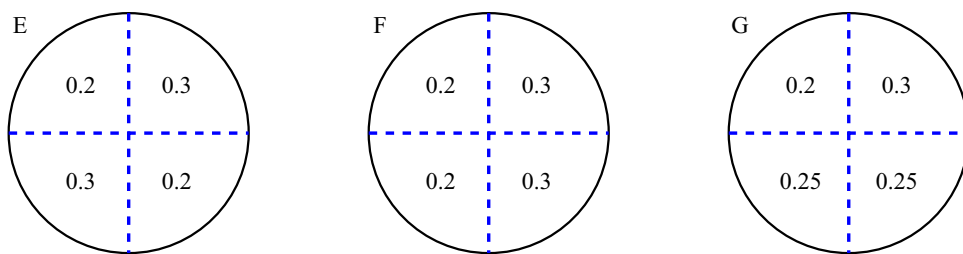
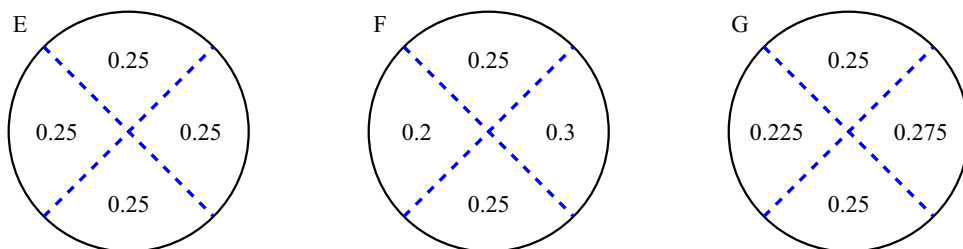**Fig. 3** Signal probabilities for model **EFG**

Fig. 4 Signal probabilities for model **EFG**, alternative description

measurements are to some extent arbitrary, and cannot be relevant for selective explanation.

To make the objection clearer, consider one of Pfeifer's models (Pfeifer 2006, p. 325). Suppose there are three types of bee, **E**, **F,** and **G**.[6] They each need to signal the location of food, and each signal indicates a direction around a circle centered on the nest. Further, they differ in the probabilities of signaling in each direction, as per Fig. 3.

Described this way, each type transmits a certain amount of information on average:

– **E**: 1.971 bits/signal
– **F**: 1.971 bits/signal
– **G**: 1.985 bits/signal

We see that **G** transmits more than **E** and **F** on average. However, if response behavior is distributed uniformly about the segments, we can individuate signals by cardinal points instead, as per Fig. 4.

Described this alternative way, the information measurements are:

– **E**: 2 bits/signal
– **F**: 1.985 bits/signal
– **G**: 1.996 bits/signal[7]

Notice that the order of greater/lesser information transmission has changed. Now **E** transmits more than **G** types,

followed by **F** types last. The subsequent argument is fairly obvious. If signal individuation is down to the modeler's discretion, and information measures are sensitive to individuation, then information transmission is model-relative.

Though there are many ways for depictions of the world to be relativized to the model used to express them, and not all render hypotheses inconsequential, the kind of relativity implied by this objection is problematic. A selective hypothesis is supposed to pick out a property and state why it persists in the lineage. In this case, the property is the quantity of information contained within signals. If that quantity cannot be said to be greater or lesser except relative to the model chosen to represent the signal, we need independent grounds on which to choose one model over another. Pfeifer claims we have no such grounds since no carving of the space is privileged (Pfeifer 2006, p. 325, col. 1): "if different ways of carving signals result in different orderings of the measurements, then it is unclear how the measure could be used in explaining how information transmission is selected for." As a result, no selective hypothesis adverting to information transmission is recommended over any other.

Response to Objection 1

There are two ways to read the objection. I will argue it fails in either case. On the first reading, it turns out recarving the space does not alter the informational measurement. The objection fails on mathematical grounds. On the second reading, recarving the space alters the measurement as required, but some carvings will more accurately match the facts than others. Selective hypotheses adverting to informational measurements are then verifiable, establishing the required link between model and reality. The objection fails on both readings. Moreover, the dilemma is not accidental. It is a reflection of Pfeifer's assumption that signals are discrete

---

[6] The original example mentioned ants. Here and below I substitute bees without loss of generality. I retain the alphabetical labeling of types for ease of comparison with Pfeifer's article. Space precludes discussion of model **ABCD**.

[7] A typo in the published version of Pfeifer (2006, p. 325) erroneously cites this value as 1.96 bits/signal.

Model EFG (type E): first reading
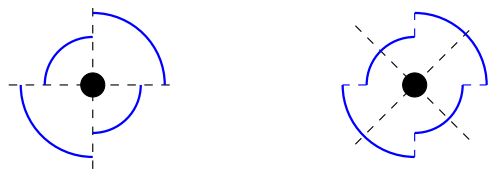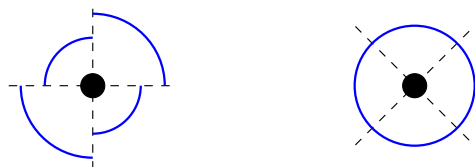

Model EFG (type E): second reading

Fig. 5 Two ways of understanding objection 1. The first entails the same informational results for both carvings. The objection is then incorrect on mathematical grounds. The second entails different informational results, but different falsifiable assumptions. The objection fails because the models' assumptions are verifiable, not arbitrary in the manner required for the objection to go through

and indicate direction symbolically. In fact, signals in this system are continuous and indicate direction iconically.

Consider the first reading. The objection rests on statistical differences between the two carvings. However, the second partitioning does not seem to have the entropy assumed by the second informational calculation. If the distribution of insects across each partition in the first carving is uniform, the distributions in the second will *not* be uniform. To see this, consider Fig. 5 (top row). Solid curved lines (blue in online version) represent insects observed in each segment as a proportion of total insects observed. These densities would not change even after the partitions are shifted, so the new distributions are nonuniform lines with a step-change in the middle of each partition. The information rate derived from such stepped distributions is the same as that calculated from the original partitioning. As a result, the proposed calculation is incorrect and does not have the consequence that information measurements are model-relative.

There is another way to read the objection, however. Suppose the scientists were faced with a choice of partition to use while conducting the experiment. They can set up their four detectors in either of the two configurations depicted by model **EFG**. Due to restrictions in the precision of the detectors, they are forced to assume a uniform distribution within each segment in either case (Fig. 5, bottom row). Given these conditions, it is true that a scientist who opts for the first configuration would derive different informational measurements than one who chooses the second. Different selective hypotheses result, but one hypothesis will be better.

One carving entails more accurate modeling assumptions, since the regions it describes in fact have a more uniform distribution of insects passing through. Alternatively, both carvings could be equally wrong. This is importantly different from being arbitrary, since there will be some other carving that is better than both. In any event, a more precise detection method would capture more fine-grained insect movements, giving rise to more accurate informational calculations.

The objection fails on both readings, a dilemma that brings out a key difference between model **EFG** and the model used by Haldane and Spurway. The latter employed a discrete approximation of a continuous model, using an equation of continuous rather than discrete entropy. Although the statistic they plugged into the equation was derived from an experiment using discrete partitions, it is a different way of modeling the situation than that typified by **EFG**. Discrete approximations to continuous models become more accurate as they become more fine-grained (Schürch and Ratnieks 2015, Fig. 1, p. 3). If the insects of model **EFG** were observed again using detectors capable of distinguishing 8, 16, or 32 partitions, the resulting information measurement would become more precise. It would not increase indefinitely, as per Pfeifer's approach.

To reiterate, Pfeifer introduces a situation in which signals are discrete and represent directions symbolically. Her objection is then that we cannot possibly come to know how symbols and directions correspond, and any choice destroys the link between model and reality. In contrast, the biologists make the reasonable and widely accepted assumption that bee signals are continuous and represent directions iconically. Once a continuous model is employed to represent a signaling system, discrete partitions can be overlaid to retrieve data. The more fine-grained these partitions, the more accurately the data captures the continuous model lying underneath.

A different but equally mistaken conclusion might be drawn from the equation used to convert statistical into informational data. To measure the accuracy of the waggle dance, Haldane and Spurway took a circle centered on the nest and measured the proportion of that circle covered by insects that had received a signal (recall Fig. 2 and Eq. 1). A smaller circle proportion covered by food-seeking receivers means more accurate communication, hence a greater amount of information transferred. Pfeifer objects to the appearance of the number 360 in Eq. (1):

> It is assumed in both cases that there are 360 possible messages about direction corresponding to the 360 degrees surrounding the nests of the ants or bees [but] it is not clear why they should be divided into 360 different possible signals, as opposed to 180, 720, or some other number. (Pfeifer 2006, p. 342)
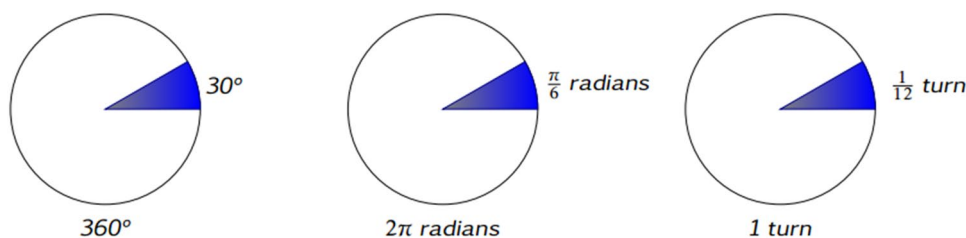
Fig. 6 The proportion of a circle covered by receiver insects can be calculated using any units, whether degrees, radians, or as a fraction of a turn. In Eq. (2), σ (Greek letter sigma) represents the angle of the shaded segment. Hence the first logarithm in Eq. (1) represents the length of an arc as a proportion of the circumference of the circle, which is not relative to the conventional measure of 360 degrees

It is easy to see the problem if the objection hits its mark. The division of circles into 360 equal segments is a human convention. If it is used as the basis of an information calculation, any derived quantities must be arbitrary as well. Conventional measurements cannot enter into explanations that pertain to natural selection, since nature is blind to our conventions. Consequently the use of 360 in Eq. (1) is inappropriate. Again, citing Pfeifer, Sarkar reiterates the objection (Sarkar 2013, p. 196).

However, the interpretation of Eq. (1) that underpins this argument is incorrect. The number 360 is employed only as a convenience for the reader. The equations in which it appears are designed to calculate the proportion of the circle covered by outgoing insects. Proportion is blind to the units used to calculate it since dimensions cancel in the division. Indeed, Haldane and Spurway first represent the full angle in radians as 2π before converting it to 360 degrees for expository clarity (Haldane and Spurway 1954, p. 255). Unfortunately they use the same symbol, σ, in both cases, making it less than obvious that the units have changed from radians to degrees. Equation (2) and Fig. 6 demonstrate the equivalence of these different measures.

$$\text{Circle proportion} = \frac{\sigma_{\text{degrees}}}{360} = \frac{\sigma_{\text{radians}}}{2\pi} = \frac{\sigma_{\text{turn}}}{\tau} \qquad (2)$$

As a result the use of the number 360 does not indicate an assumption of 360 messages. It is true, however, that the *precision* of the calculation increases as the precision of σ increases. Pfeifer raises a related complaint that the statistical data available to Haldane and Spurway are given in 15 degree increments, which is too coarse-grained to provide a reliable measure. However, the biologists take this into account, and the error margins they give are consistent with their estimations. As per Schürch and Ratnieks (2015, p. 3), more precise measurements would lead to more accurate calculations. In sum, this objection is not sustained.

Overall then, the accusation of arbitrariness toward coordinate systems used to measure the directional component of the bee dance cannot be upheld. Continuous models are available for use where appropriate. Haldane and Spurway employed such a model. Though the literature on sender-receiver models emphasizes discrete signal sets, this is not a necessary feature of the framework. Plausibly, many signals in nature can be profitably investigated with continuous models. Discrete models are often best construed as approximations to the underlying continuous representation. In the next section I deal with a second major objection to the use of information theory in the study of biological communication.

Objection 2: Precision and Accuracy

In this section I show how quantifying information in signals is not just a case of quantifying the precision of receiver behavior. One route to skepticism of the utility of information theory in biology is based on the claim that behavioral precision does not entail accuracy, hence precision alone entails nothing about fitness. I show that information measurements, far from being solely a measure of precision, are a quantification of accuracy. Accuracy, which in this case translates into foraging efficiency, is clearly relevant for fitness.

The objection, again due to Pfeifer, runs as follows. The biologists used statistical data in their calculations, and the statistic they employed is the spatial distribution of signal receivers. A greater proportion of receivers clustered closer to the target entails a greater quantity of information transmitted. By measuring the difference between this more "focused" distribution, and the otherwise random distribution of receivers around the circle, the biologists quantify information. Pfeifer complains that mere increased precision of receiver behavior might not be selectively relevant. Suppose bees clustered closely together in a region that lacked food. Their behavior would be focused in a statistical sense, yet would not contribute to fitness. Behavior can be highly specific yet selectively neutral, or even detrimental. To support the objection Pfeifer presents model **HIJ**.

Suppose there are three kinds of bee, **H**, **I**, and **J**, each of which can signal to the east and west and do so with equal probability. Suppose **H** types always go east when signaled east and west when signaled west, but **I** types respond

improperly by going west when signaled east and east when signaled west. In addition, **J** types go east with probability 0.7 when signaled east, otherwise west, and vice versa. The problem highlighted by this model is that **H** and **I** types transmit the same amount of information, namely 1 bit/signal, whereas **J** types transmit much less, around 0.119 bits/signal. Despite this, we would expect **J** types to perform better than **I** types because they successfully obtain food more often on average. In consequence, it is unclear how information transmission can have relevance for fitness.

Response to Objection 2

We can respond to this objection by noting that the increased "focus" of the posterior distribution is centered on a target that has selective relevance, such as food. Wilson (1962) confirms that this assumption is encapsulated by the model. The parameter used to measure information, σ_d, is directly relevant for accuracy:

> Note that σ_d refers to dispersion of following workers with reference to the actual target and hence is a direct measure of the accuracy (content) of information transmission. The honeybee data of von Frisch and his co-workers are of this nature. (Wilson 1962, Appendix, p. 157, col. 1)

Insofar as receiver bees have a greater probability of finding food after having watched a dance, information is transmitted by dancing bees. It is this notion of transmission that constitutes grounds for application of the information concept, and related mathematical tools, in studies of biological signaling.

What about model **HIJ**? It seems to show that biologists might inadvertently attribute a high rate of information transmission to a colony that fails to put its signals to good use. As a claim about methodology, this is implausible. If it is assumed that signals used by type **I** are indeed food signals, it would be mysterious why receivers behave so aberrantly. Where the colony's goal is finding food, signals generally increase the probability of doing so. When they do not, they are selected against. **I**-type signals could hardly be regarded as signals about food since the probability of finding food is not raised when they are sent.

Like any piece of behavior, a signal cannot be selected for if it does not contribute to fitness. By hypothesis **I**-type signals do not contribute to fitness, therefore they cannot be selected for. As a result they would not contribute to a behavioral ecologist's selective hypothesis; they would be simply anomalous.

To reiterate, the biologists' models embody a crucial assumption: that the target has selective relevance. In this section, two mathematical objections have been overturned. I now turn to broader considerations about the use of informational concepts in animal communication theory, paying particular attention to the special nature of the honeybee waggle dance and how the arguments presented here are affected by weakening some of the assumptions required to model it.

Wider Considerations of the Use of Informational Concepts and Measurements

As a result of the foregoing interpretation of honeybee signaling, we are in a position to consider several of ORR's challenges in a new light. As this article is primarily defensive, the alternative influence-based definition is not critically assessed, nor do I consider systems in which animals possess mental representations that may be used to produce or interpret signals. I deal only with information in the sense Kalkman (2017, p. 1127) describes as an "ultimate explanatory construct." The target of explanation is fitness-improving behavior, including coordination of two or more agents. In proceeding via minimal models of communication, nothing need be assumed about cognitive sophistication. To be sure, proponents of information in such signals as primate alarm calls must face other challenges from ORR. I remain silent on those aspects of the debate.

Defining Information and Encoding

The most pressing challenge to the optimist about information is the lack of a clear definition of that concept in behavioral ecology. I introduce the challenge before responding to it. ORR point out that many authors use the quasi-technical term in an unconstrained manner (Rendall et al. 2009, Table 2; Rendall and Owren 2013, Table 6.1). As a result those authors often switch between different concepts masquerading under the same label, or gesture at unsubstantiated explanations. In particular, authors often invoke "information" in a technical sense that quantifies correlations between events. This *Shannon information* cannot be central to definitions of signaling, however, because it can describe any correlated events, biological or otherwise.

This distinction between the measurement of correlations and a richer, semantic notion of information lies at the heart of this debate and wider issues surrounding biological information. The relevant *Stanford Encyclopedia of Philosophy* article as of 2018 (Godfrey-Smith and Sterelny 2016) organizes its exposition around two distinct approaches: "Shannon's concept of information" and "Teleosemantic and other richer concepts." Sarkar (2013) distinguishes Shannon information (he calls it MCT-information) from semantic information, and repeats ORR's contention that theorists often conflate the two. The distinction has also been flagged

as between "syntactic" and "semantic" information, the former being agnostic about the meaning of signals or symbols whose transmission it quantifies (Morton and Coss 2013, pp. 211, 229). The consensus is clear: Shannon information has no import for semantic information, and theorists do wrong to conflate the two. Since invocations of information are inconsistent, and nobody has provided a resolute definition, it seems better to dispense with the notion.

A similar fate befalls the concept of encoding. Signals are sometimes said to carry information in a form that a receiver must decode in order to obtain. But it is rarely made explicit what is meant by an "encoded message" in a biological setting. By failing to specify what is being invoked, theorists draw on a "vague, elastic and insubstantial" concept (Owren et al. 2010, p. 758) that cannot do the explanatory work required. As with "information," then, talk of "encoding" does more harm than good. Both can be discarded without losing explanatory power. Rendall and Owren (2013, pp. 171–172) make an even stronger claim: "Ultimately, notions such as information and coding cannot be cashed out in terms of standard concepts used in biological and evolutionary theory."

One thing worth mentioning here is that ORR's wider aim, to critique anthropocentrism and linguistic metaphors in animal signaling theory, is laudable. Animal communication should indeed be understood on its own terms, and should not be treated as a pale imitation of human natural language. It is interesting, however, that ORR consider the use of information theory to fall into this anthropomorphic paradigm, since several theorists present the story precisely the other way round. The "conduit metaphor" that ORR discuss is a term introduced by Reddy (1979) in part to highlight the limitations of modeling human communication as an MCT channel. Reznikova (2017) introduces information-theoretic methods in studies of communication between ants, in an attempt to get away from language-inspired metaphors and methodologies. Harms (2004) promotes formal methods for capturing meaning without recourse to translation, since translating animal signals into human language can only mislead us about the nature of meaning. All these theorists note the disparity between MCT and human language. Below, I continue work in that tradition, arguing that information theory is apt for generalization to animal communication theory precisely because it captures something far more general than human engineering constructs. It should become clear that information theory, rightly interpreted, is no more anthropocentric than any other branch of mathematics.

Claims about the irrelevance of information measures for "semantic information" are deeply mistaken, and are discussed in the next subsection. I want first to show that both information and encoding are clearly defined in the honeybee case. As discussed at length above, the measurement

of transmission rate captures the accuracy of receiver bees, hence the efficiency of foraging behavior. Haldane and Spurway's insight was to define an entropy over the space of relevant behavioral outcomes such that when entropy is reduced by a communicative act, the measure of information rate is simply a measure of successful functional performance. This is an easy measurement to perform in the honeybee case because foraging efficiency is linked to spatial accuracy, and spatial accuracy is comparatively easy to ascertain.

So information can be defined, at least in this case, in terms of functional performance, rendering it both tractable and relevant for selection. Consider now the concept of encoding. After a period of controversy in the middle of the 20th century, it became widely accepted that the different components of the bee dance correspond to different spatial relations between the hive and the indicated food source. These aspects of the dance vary with those spatial relations in a principled way. If they did not, receivers would be unable to reliably exploit the relevant food source. Although there are several other factors affecting transmission rate, the correspondence between spatial relations and dance features—what I am here calling the code—clearly affects it too. Consider the round dance, an alternative to the waggle dance used when food sources are very close. This dance has no directional component, meaning that on average its transmission rate would be lower than the waggle dance. Greater articulation affords greater transmission rate, which presumably contributes to the explanation why the waggle dance is articulated into different components. In general, then, a code is a set of correspondences between signal and world, designed so that receiver behavior covaries aptly with the state of the world.[8]

In a section below I discuss prospects for extending these definitions to cases of animal signaling not typically characterized by common interest. Having demonstrated these definitions, I must now clear up a deep misconception regarding the use of information theory in biology.

Information and Meaning

One of the conceptual advances that heralded the development of communication theory was the recognition that meaning could be abstracted away from symbol sequences. Common interpretation of a symbol sequence means that all that need be passed in order to effectively pass meaning is

---

[8] This definition has pedigree from at least two sources. First, Millikan's definition of mapping rules between signal and world sustains the requirement of codesign and takes as a canonical example the articulation of the waggle dance (Millikan 1984, p. 107). Second, Skyrms's account of the evolutionary emergence of conventional meaning outlines correspondences, resulting from coadaptation, between signals and the behaviors they cause (Skyrms 2010, §§3–5).

the sequence itself. Subsequently, measures of information were defined in terms of the passability of the sequence, prohibiting mention of any meaning it might be associated with. This hallmark of the basic model of MCT is expressed by Weaver's oft-quoted warning:

> The word *information*, in this theory, is used in a special sense that must not be confused with its ordinary usage. In particular, *information* must not be confused with meaning. In fact, two messages, one of which is heavily loaded with meaning and the other of which is pure nonsense, can be exactly equivalent, from the present viewpoint, as regards information. (Shannon and Weaver 1949, p. 8; emphasis in original)

Through Bar-Hillel and Carnap (1953), Dretske (1983), Dennett (1983), and Krebs and Dawkins (1984),[9] that warning found its way into the animal signaling literature. ORR cite it unmodified (Rendall et al. 2009, p. 240, col. 2; Owren et al. 2010, p. 761), and Sarkar (2013, p. 193) reaffirms the sentiment. As mentioned above, the canonical view in the philosophy of biology is that mathematical concepts from information theory must be different from whatever is meant by the "meaning of" or "semantic information carried by" a biological signal.[10]

The central MCT model abstracts meaning away as described above. But what was not recognized until much more recently is that it smuggles in a different kind of meaning. Codesigned signalers and receivers passing messages across a medium are apt for a straightforward teleosemantic analysis. The encoded message of the MCT model fits this description perfectly. Regardless of the meaning of the original symbol string, encoded messages possess the kind of meaning Godfrey-Smith (2013) calls "subpersonal content."[11] Encoded messages stand for the source message from which they were derived. The code is shared between

transmitter and receiver as a consequence of codesign. Electrical pulses in telegraph wires are paradigmatic examples of signals with subpersonal content. It is crucial to note that this holds regardless of whether the symbol strings they encode have any further meaning. Simply being encoded and decoded by a codesigned signaler-receiver pair is enough to count as a contentful signal. In the special case of MCT, what signals signify is another string of symbols.

The consequences of this perspective will be crucial for understanding the relationship between MCT and biological signals, including animal communication. First, notice that what is quantified by transmission rate is the efficiency of joint function. The only goal of the central MCT model is to reconstruct a source message at the target. This is true whether the signal is a crucial battle order or a nonsense string compiled by an infant. Its success in this endeavor is what is quantified by transmission rate.[12] Maximum possible success is quantified by channel capacity. What is almost universally ignored is that message reconstruction is not the only joint goal whose efficiency can be quantified in this way. Signals need not stand for and produce strings of symbols, as they inevitably do in MCT.[13] They can stand for states of affairs and produce behavior directly, without further symbolic intermediaries. The honeybee waggle dance is a canonical example.

The perspective of subpersonal content affords both a naturalistic approach to meaning as well as insight into the special nature of the MCT model. There is one further consequence worth mentioning that highlights an important but tacit assumption lying behind at least one of ORR's objections to information-talk (see below). In many discussions of this and related topics, "information" is taken to imply an exclusively indicative correspondence between signal and world. This can be seen whenever it is cashed out in terms of the probability of an outcome, correlations between events, a relation of standing-for, or what a receiver could infer from the signal. What has come out of the literature on subpersonal content is another kind of correspondence between a signal and the world. Signals often bear an instructive element, being supposed to bring about some state of affairs. This suggests an interpretation in terms of the favorability of an outcome, or of what a signaler can control. Signals may be more or less instructive or informative; in the simplest systems characterized by full common interest they possess both aspects to the same extent.[14] To continue our earlier

---

[9] The closest I can find to the traditional irrelevance claim is on p. 395: "Measurements of Shannon information do not necessarily reveal anything about semantic information, although they often do."

[10] There is at least one prominent school of dissent to this orthodoxy, in the shape of the sender-receiver paradigm headed by Brian Skyrms (2010). Though Skyrms accepts that the total quantity of information in any given signal is silent on its content, he proposes a definition of information content that makes use of informational measurements (Skyrms 2010, §3). Importantly, it does not cut between signals and cues. Like ORR's definition of Shannon information, it is defined in terms of probabilities, not function. In what follows I leave the Skyrms account to one side. Though it is useful and innovative, there is a much more direct way to demonstrate the relationship between information in MCT and meaning in biological signals.

[11] "Subpersonal" is applied to brain states that are not assumed to play a role in conscious thought. I use it here as a modifier to "content" that makes no assumptions further than codesign of signalers and receivers. It is therefore applicable to subcognitive and noncognitive systems in the manner of Shea (2007) and Millikan (2013).

[12] Measured in bits per symbol, not per second.

[13] Oliver Lean has made the same point (Lean 2016, pp. 239–240).

[14] Signals with this dual character were labeled "neutral" by Lewis (1969), "pushmi-pullyu" by Millikan (1995) and "primitive" by Harms (2004). The latter is the preferred term. Incidentally, it might be thought that the instruction/information distinction is just another way of describing the influence/information distinction. Unfortunately, the latter is far more ambiguous.

example, an encoded message in the central MCT model equally informs the receiver about the source message and instructs it which target message to produce. Because the literature on information theory places so much emphasis on inference, this instructive aspect of signals is often missed.

In sum, the traditional distinction between semantic information and formal measures of transmission in communication channels is misguided. By neglecting the special nature of the MCT model, scholars have improperly adopted Weaver's irrelevance claim in domains where it does not belong. I now turn to a more pressing objection to the use of information theory in behavioral ecology.

Information, Costs and Benefits

Sarkar (2013, pp. 200–201) offers a further challenge to this construal of information. Although animals' evolutionary interests can be furthered by certain behavioral strategies, and although these strategies sometimes require coordination with other animals, describing and explaining optimal behavior never requires information-theoretic formalism. Speaking of the honeybee waggle dance, he claims "we can quantify the loss of optimality (e.g., as a function of the departure from optimality of resources gathered relative to foraging effort) with no recourse to 'information' " (Sarkar 2013, p. 201).

On the surface this claim seems plausible. Recall Haldane and Spurway took one piece of statistical data, the standard deviation of the Gaussian distribution, and converted it into a measurement of transmission rate. It is not clear what that latter calculation adds to an evolutionary analysis, since describing and explaining optimality seem to require only tallying costs, benefits, and probabilities. As it stands, however, Sarkar's claim is unsubstantiated. As Sarkar rightly notes, many previous attempts to quantify aspects of animal signaling systems measure only the entropy of signaling repertoires, if they measure anything at all (Sarkar 2013, §4). The best chance of extending the concepts of MCT to animal signaling theory might be to explain the form of a signal by reference to the signalers' need to coordinate. As well as correspondences between signal and world, properties such as redundancy and noise contribute to this analysis. A signal is a piece of behavior, and like behavior more generally it can be optimized. The mathematics describing the optimal form of a signal, where signaler and receiver are codesigned, is communication theory.[15]

In their more tempered approach to redefining animal communication, Krebs and Dawkins (1984, p. 396) suggest information theory might be best suited for an analysis of signaling economics rather than measuring bit rate. Striking a balance between energy expended on signaling and ensuring sufficient coordination falls neatly into the optimization paradigm already well deployed for individual behaviors. It does not seem implausible that this can be extended to joint behaviors, nor that some appropriate generalization of communication theory can play a role in the mathematization of this extension.

On reflection, at least part of the reason why it is difficult to draw insight from Haldane and Spurway's measure is that it is difficult or impossible to measure costs and benefits, hence difficult to estimate the parameters of "waggle dance economics." The second and third sections, above, are dedicated to establishing that Haldane and Spurway measured something real—the transmission rate of the waggle dance of von Frisch's bees—and that their results were not arbitrary, contrary to recent objections. To go beyond this, we would need to add hypotheses about the selective environment of the honeybee and the context in which signals were useful (Preece and Beekman 2014). We would need to target specific hypotheses about the process by which the primordial cue became the contemporary signal. This would likely involve comparative work on dialects between different species and subspecies (Beekman et al. 2015). None of this threatens the literal truth of the information rate measure, though it may make it difficult to integrate it with other empirical work. In particular, I concur with Sarkar that we have reason to be agnostic about the significance of Wilson's comparison between the signaling rates of honeybees and fire ants (Wilson 1962). I am skeptical that this comparison could be meaningful without significant assumptions about the foraging ecology and evolutionary history of both bees and ants. Comparing transmission rates may be no more valid than comparing costs and benefits between very different species. For closely related species this might be a useful tactic, but in most cases, such comparisons are meaningless.

Extending the Method to Other Cases

Several reasons counted in favor of using the bee dance as a case study. Through decades of research there is plenty of data available on bee dances and their effect on compatriot bees. Research continues, and any empirical hypotheses produced by the theoretical work advocated here could be tested on several subspecies. In addition, since bee sociality is relatively recent and quite diverse (Danforth 2007), links between communication and cooperation could be investigated in depth using the bee family. With regard to honeybees, strong eusociality entails a negligible threat of free

---

[15] The distinction between code-as-correspondence and code-as-redundancy mirrors a well-known duality in communication theory between data compression ("source coding") and data transmission ("channel coding"). Speculatively, the selective pressures on animal communication may be skewed far more towards solving the latter problem. Hailman's book (Hailman 2008) and Wiley's essays (Wiley 1983, 2013) are examples.

riding or defection. It is reasonable to assume signaling is fully cooperative, avoiding complex problems of cross-purpose and deception. But the question must be faced of what becomes of this methodology in cases where few or none of these assumptions hold. In this section I consider what we have learned so far in the light of more pressing objections.

Explanations of Signaling Behavior Must Not Ignore Physiological Constraints

As a consequence of the abstraction required for informational description, the question of behavioral and developmental constraints has often been ignored. Animals do not have great freedom of action, and natural selection does not have free rein in shaping their behavior. That many purported instances of signaling can be shown to be distinctly manipulative is a consequence of this fact (Rendall et al. 2009, p. 237; Owren et al. 2010, pp. 766–767). Consider two populations of signalers and receivers without perfectly aligned interests, such as males and females of a single species. Suppose, as a general rule, overt behavior can adapt faster than the cognitive architecture underlying perception. Then receivers will often lag behind in the perceptual arms race. Signalers will be quick, on an evolutionary scale, to exploit perceptual biases while receivers will be slow to rectify them. Receiver responses are then determined by trade-offs between the need to respond to relevant perceptual stimuli without being duped too harshly.[16]

This story of trade-offs goes beyond signals that are "about" something, such as male quality, and includes behaviors that cannot be evaluated in terms of honesty. ORR point out that many examples of what we typically think of as signals are not (or only degenerately) contentful, because their function is tied specifically to receiver affect or attention. For example, males of some fish species display coloration similar to that of their female receivers' prey, the point being to catch a female's attention (Owren et al. 2010, p. 767). Importantly, being better at attention-grabbing need not correlate with quality. The reason for this male adaptation is not a corresponding female adaptation; it is a prior female adaptation for an entirely different function.

There seem to be three intertwined issues here. First, traits and behaviors designed as attention-grabbing appear not to be contentful. Second, physiological constraints play a larger role in shaping the form of the signal than is usually afforded by information optimists. Finally, male coloration is not a consequence of coadaptation. I deal with the third

problem in the next subsection, and will speak now to the problems of content and constraint.

In the example, fish coloration is not contentful, because it does not correlate with anything. Note, however, that the assumption lying behind this objection is that informational language entails that signals *indicate* a state of the world. As argued above, signals can also contain instructions how to act. This could include an instruction to "pay attention," opening the channel for further signaling. Of course, this still requires coadaptation, and by assumption the fish in the example do not meet this condition. I deal with this wider problem in the next subsection, but the point to take away here is that coadapted instructive signals need not have the familiar indicative content usually assumed by paradigms like costly signaling theory.

Consider now physiological constraints. The informational approach seems to assume the form of a signal should be somewhat divorced from its content, so that it can be freely optimized for reliable transmission. But signaling behaviors whose form is primarily attributable to physiological and developmental constraints are not apt for this explanation. It looks like we need an *evolutionary* communication theory, one that considers optimal joint behavior from the perspective of restricted design capabilities. Models of communication from an evolutionary perspective would afford much less freedom of choice of encoding. In particular, the code—the form of the signal—that develops through coevolution will be simultaneously constrained by competing adaptive needs of both the signaler and receiver. Crucially, it could still be the case that coadaptation plays a role in explaining signal form, even though several different kinds of constraint play a role too. In the case of the waggle dance, certain physiological constraints have been proposed to explain a systematic error in dance performance (Preece and Beekman 2014). This is the kind of hypothesis that could be integrated with a formal approach to signaling theory, provided that other concerns detailed in this article are addressed.

ORR have clearly presented important considerations for evolutionary perspectives on communication. All behavior is subject to physiological and developmental constraints. Natural selection does not give rise to perfect forms. Evolutionary models tend to make assumptions trading realism for tractability. We must at least find some way of incorporating proximate constraints into these investigations.

Explanations of Signaling Behavior Must Not Ignore Divergent Interests

Continuing the example from the previous section, there were three conditions that the fish signals failed to meet: being contentful, being designed with few or no constraints, and being coadapted. I agree we must accommodate a

---

[16] Though see Bergstrom and Lachmann (2003) for conditions under which the slower-evolving organism enjoys the benefit, a phenomenon they call the Red King effect.

broadened perspective on the first two conditions, but see no reason why a mathematical toolkit provided by a generalized communication theory should not play a part. Now consider how the informational approach should deal with "signals" that are not the result of coadaptation.

One option here is to reject that those traits are signals. They seem instead to be the logical converse of cues: correlated behavior that occurs for the signaler's benefit rather than the receiver's. The traditional distinction becomes a tripartite carving between cues, signals, and manipulations. ORR include the latter two categories in their redefinition of signaling. I want to highlight that it is at least possible to separate them, though I accept that the attractive symmetry of the tripartite distinction is unlikely to be representative of our asymmetric world.

Just as with cues, manipulations blend into signals when previously unidirectional influence becomes coadaptive. These three categories are rough coverings of many possibilities. Evolutionary game theory is explicitly designed to study these situations. In particular, work inspired by Skyrms's sender-receiver framework has explored the limits of communication in noncooperative settings (Wagner 2012, 2015; Martínez and Godfrey-Smith 2016), as well as different rates of evolution (Brusse and Bruner 2017). Further, making this distinction will likely help ORR achieve at least one of their goals. They aim to stop theorists searching for the wrong explanations of signaling behavior. By adopting a third category, they can vividly point out that the majority of cases are closer to manipulations than coadapted signals. This allows us to distinguish cooperative situations, typified by the social insects, from those with multiple conflicting evolutionary forces.

All these problems were part of the motivation of Dawkins and Krebs's original suggested redefinition of signaling:

> To summarize the point of view we are adopting: as an inevitable byproduct of the fact that animals are selected to respond to their environment in ways that are on average beneficial to themselves, other animals can be selected to subvert this responsiveness for their own benefit. This is communication. (Dawkins and Krebs 1978, p. 285)

The authors explicitly conflate manipulations and coadapted signals. By contrast, the informational approach highlights an important difference between them. Just as parasites must walk a fine line between exploiting and destroying their hosts, so signalers must find a balance between reliably manipulating receivers and driving them out of existence. Basic models proceed on the assumption that signaling can only exist when some mechanism helps maintain equilibrium, preventing one or both parties going extinct. One obvious example of such a mechanism is when the same behavior brings benefit to both signaler and receiver. In these cases there is an element of common interest, hence there will be some level of coadaptation. The signal is then partly amenable to analysis in terms of information and encoding.

Above, we agreed theoretical and modeling work should continue to expand its horizons regarding manipulation and constraint. Now we can better see the motivation behind costly signaling theory. One party in an interaction characterized by partial common interest may be at risk of extinction, but which one (and why) will be an empirical question in each case. The mechanisms evolution has thrown up that manage to prevent extinction, thus prolonging signaling behavior, may be many and varied. Part of the work of theory building is to find a taxonomy of such mechanisms and outline their symptoms, which can then act as diagnostics for field workers to employ (see, for example, Hurd and Enquist (2005)). Dawkins and Krebs promote the view, later picked up by ORR, that in the vast majority of cases we will find general-purpose receiver perceptual mechanisms being exploited by special-purpose signaling behavior. So be it: theoretical work will uncover a wider class of forces than those manifested in nature.

All of this is apt—none of it threatens our account of information. Evolutionary game theory is designed to explore interactions of divergent interests, and it shades into communication theory when the interests are common and there are fewer physiological constraints. Molecular biology is working in this direction (Nakano et al. 2013; Iglesias 2016), sharing with social insect studies the freedom to ignore divergent interests. But weakening the assumption of common interest does not mean we cannot use mathematics to describe it. It means that the mathematics gets more general and its interpretation perhaps more obscure. It also means there will be mechanisms other than "honesty" maintaining signaling interactions, and that we should not necessarily look for static equilibrium behavior, but ongoing arms races. Sarkar believes that none of this mathematics will have anything to do with information theory. Throughout this article, I have tried to show that is an overly narrow conception of the links between communication theory and game theory.

Conclusion

In this article I have focused on the honeybee waggle dance, demonstrating how directional information in the dance is measured. The quantity of information derived is a measure of the accuracy of receiver bees. Increased accuracy entails increased efficiency of finding food over a blind search strategy. Clearly these considerations have relevance for the fitness of foragers, and the colony as a whole. The objection that information measurements are arbitrary, or have no relevance for fitness, cannot be sustained.

Several obstacles stand in the way of attempts to generalize this approach to other cases of animal signaling. For one thing, signal form is not always explicable by reference to coadaptation. The definition of encoding advocated here will not be available. Further, it is still unclear how to unify informational measurements with fitness cost. It seems likely modeling work will continue even while it remains difficult to test with empirical examples.

However, genuine concerns should be distinguished from those raised due to misunderstandings of communication theory. Claims about the undefinability of information, or its irrelevance for semantic meaning, or the arbitrariness of its measurement have been criticized here. In particular, claims about the irrelevance of communication theory for semantic information are typically confused. The central model of MCT employs symbols that stand for symbols, which is why the meaning of one of those sets of symbols is irrelevant for quantification. But it is the function of the other set, the meta-symbols, that is quantified by transmission rate, and whose design is optimized for better performance. Future work will seek to characterize far more vividly the relationship between evolutionary game theory and an empirically informed evolutionary communication theory.

References

Bar-Hillel Y, Carnap R (1953) Semantic information. Br J Philos Sci 4(14):147–157

Beekman M, Gloag RS, Even N et al (2008) Dance precision of Apis florea—clues to the evolution of the honeybee dance language? Behav Ecol Sociobiol 62(8):1259–1265

Beekman M, Makinson JC, Couvillon MJ et al (2015) Honeybee linguistics—a comparative analysis of the waggle dance among species of Apis. Front Ecol Evol 3:11

Bergstrom CT, Lachmann M (2003) The red king effect: when the slowest runner wins the coevolutionary race. Proc Natl Acad Sci USA 100(2):593–598

Biesmeijer JC, Seeley TD (2005) The use of waggle dance information by honey bees throughout their foraging careers. Behav Ecol Sociobiol 59(1):133–142

Brusse C, Bruner J (2017) Responsiveness and robustness in the David Lewis signaling game. Philos Sci 84(5):1068–1079

Chittka L (2004) Dances as windows into insect perception. PLoS Biol 2(7):898–900

Danforth B (2007) Bees. Curr Biol 17(5):R156–R161

Dawkins R, Krebs JR (1978) Animal signals: information or manipulation? In: Krebs JR, Davies NB (eds) Behavioural ecology: an evolutionary approach, 1st edn. Wiley, Hoboken, pp 282–309

Dennett DC (1983) Intentional systems in cognitive ethology: the "Panglossian paradigm" defended. Behav Brain Sci 6(3):343–355

Dretske F (1983) Précis of *Knowledge and the Flow of Information*. Behav Brain Sci 6(1):55–63

Dronamraju KR (2017) Popularizing science: the life and work of JBS Haldane. Oxford University Press, New York

Godfrey-Smith P (2013) Signals, icons, and beliefs. In: Ryder D, Kingsbury J, Williford K (eds) Millikan and her critics. Wiley, Hoboken, pp 41–58

Godfrey-Smith P, Sterelny K (2016) Biological information. In: Zalta EN (ed) The Stanford encyclopedia of philosophy. Summer 2016 edition. Stanford University, Stanford

Gould JL (1975) Honey bee recruitment: the dance-language controversy. Science 189(4204):685–693

Hailman JP (2008) Coding and redundancy: man-made and animal-evolved signals. Harvard University Press, Cambridge

Haldane JBS, Spurway H (1954) A statistical analysis of communication in "Apis mellifera" and a comparison with communication in other animals. Insect Soc 1(3):247–283

Harms WF (2004) Primitive content, translation, and the emergence of meaning in animal communication. In: Oller DK, Griebel U (eds) Evolution of communication systems: a comparative approach. MIT Press, Cambridge, pp 31–48

Hurd PL, Enquist M (2005) A strategic taxonomy of biological communication. Anim Behav 70(5):1155–1170

I'Anson Price R, Grüter C (2015) Why, when and where did honey bee dance communication evolve? Front Ecol Evol 3:125

Iglesias PA (2016) The use of rate distortion theory to evaluate biological signaling pathways. IEEE Trans Mol Biol Multi-Scale Commun 2(1):31–39

Kalkman D (2017) Information, influence, and the causal-explanatory role of content in understanding receiver responses. Biol Philos 32(6):1127–1150

Krebs JR, Dawkins R (1984) Animal signals: mind-reading and manipulation. In: Krebs JR, Davies NB (eds) Behavioural ecology: an evolutionary approach, 2nd edn. Blackwell Scientific, Oxford, pp 380–402

Lean O (2016) Biological information. PhD thesis, University of Bristol, Bristol

Lewis D (1969) Convention: a philosophical study. Blackwell, Oxford

Martínez M, Godfrey-Smith P (2016) Common interest and signaling games: a dynamic analysis. Philos Sci 83(3):371–392

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

Millikan RG (1995) Pushmi–Pullyu representations. Philos Perspect 9:185–200

Millikan RG (2013) Natural information, intentional signs and animal communication. In: Stegmann UE (ed) Animal communication theory. Cambridge University Press, New York, pp 133–146

Morton ES, Coss RG (2013) Mitogenic rays and the information metaphor: transmitted information has had its day. In: Stegmann U (ed) Animal communication theory. Cambridge University Press, New York, pp 207–231

Nakano T, Eckford AW, Haraguchi T (2013) Molecular communication. Cambridge University Press, New York

Owren MJ, Rendall D, Ryan MJ (2010) Redefining animal signaling: influence versus information in communication. Biol Philos 25(5):755–780

Pfeifer J (2006) The use of information theory in biology: lessons from social insects. Biol Theor 1(3):317–330

Preece K, Beekman M (2014) Honeybee waggle dance error: adaption or constraint? Unravelling the complex dance language of honeybees. Anim Behav 94:19–26

Reading A (2011) Meaningful information, volume 1 of Springer briefs in biology. Springer, New York

Reddy MJ (1979) The conduit metaphor: a case of frame conflict in our language about language. In: Ortony A (ed) Metaphor and thought. Cambridge University Press, Cambridge, pp 254–83

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

Rendall D, Owren MJ, Ryan MJ (2009) What do animal signals mean? Anim Behav 78(2):233–240

Reznikova Z (2017) Studying animal languages without translation: an insight from ants. Springer, New York

Riley JR, Greggers U, Smith AD et al (2005) The flight paths of honeybees recruited by the waggle dance. Nature 435(7039):205–207

Ryan MJ (2013) The importance of integrative biology to sexual selection and communication. In: Stegmann U (ed) Animal communication theory. Cambridge University Press, New York, pp 233–255

Sarkar S (2013) Information in animal communication: when and why does it matter? In: Stegmann UE (ed) Animal communication theory. Cambridge University Press, New York, pp 189–205

Schürch R, Ratnieks FLW (2015) The spatial information content of the honey bee waggle dance. Behav Evol Ecol 3:22

Shannon CE (1948a) A mathematical theory of communication (part 1). Bell Syst Tech J 27(3):379–423

Shannon CE (1948b) A mathematical theory of communication (part 2). Bell Sys Tech J 27(4):623–656

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

Shea N (2007) Consumers need information: supplementing teleosemantics with an input condition. Philos Phenomenol Res 75(2):404–435

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

Stegmann UE (ed) (2013a) Animal communication theory: information and influence. Cambridge University Press, New York

Stegmann UE (2013b) On the 'transmission sense of information.' Biol Philos 28(1):141–144

von Frisch K (1948) Gelöste und ungelöste Rätsel der Bienensprache. Naturwissenschaften 35:38–43

von Frisch K (1950) Bees: their vision, chemical senses, and language. Cornell University Press, Ithaca

von Frisch K (1952) Die wechselseitigen Beziehungen und die Harmonie im Bienenstaat. Number XXXIV in Colloques Internationaux du CNRS, Paris, pp 271–292

Wagner EO (2012) Deterministic chaos and the evolution of meaning. Br J Philos Sci 63(3):547–575

Wagner EO (2015) Conventional semantic meaning in signalling games with conflicting interests. Br J Philos Sci 66(4):751–773

Wiener N (1948) Cybernetics; or, control and communication in the animal and the machine, 2nd edn. MIT Press, Cambridge

Wiley RH (1983) The evolution of communication: information and manipulation. In: Halliday T, Slater PJB (eds) Animal behaviour: communication, vol 2. Blackwell Scientific Publications, Oxford, pp 156–189

Wiley RH (2013) Communication as a transfer of information: measurement, mechanism and meaning. In: Stegmann U (ed) Animal communication theory. Cambridge University Press, New York, pp 113–132

Wilson EO (1962) Chemical communication among workers of the fire ant Solenopsis saevissima (Fr. Smith) 2. An information analysis of the odour trail. Anim Behav 10(1–2):148–158