



The relevance of communication theory for theories of representation

Stephen Francis Mann^{a,b}  (stephenfmann@gmail.com)

Abstract

Prominent views about representation share a premise: that mathematical communication theory is blind to representational content. Here I challenge that premise by rejecting two common misconceptions: that Claude Shannon said that the meanings of signals are irrelevant for communication theory (he didn't and they aren't), and that since correlational measures can't distinguish representations from natural signs, communication theory can't distinguish them either (the premise is true but the conclusion is false; no valid argument can link them).

Keywords

Communication theory · Information theory · Intentionality · Representation · Shannon information · Teleosemantics

1 Introduction

1.1 A tension

Communication theory measures the costs and benefits of representation, and describes judicious representational trade-offs. There is a popular idea that the aspects of representation the theory deals with are strictly distinct from representational content:

[Communication theory] ignores questions having to do with the *content* of signals, what *specific information* they carry, in order to describe *how much* information they carry.

Dretske (1981, p. 41), emphasis original

Shannon-Weaver theory measures the *capacity* of information-transmission and information-storage vehicles, but is mute about the

^a University of Barcelona.

^b Max Planck Institute for Evolutionary Anthropology.



contents of those channels and vehicles, which will be the topic of the still-to-be-formulated theory of semantic information.

Dennett (1983, p. 344), emphasis original

Shannon information does not capture, nor is it intended to capture, the semantic content, or meaning, of signals.

Piccinini & Scarantino (2011, p. 21)

Shannon's theory, taken in itself, is purely quantitative: it ignores any issue related to informational content.

Lombardi et al. (2015, p. 1989)

Shannon offers no analysis of the relation in virtue of which a sign carries information *about* a state of affairs (his interest was in other issues).

Neander (2017, p. 7), emphasis original

The problem with these claims – the tension this paper will try to resolve – is that they are contradicted by the words of communication theorists themselves:

[Efficiency is achieved] in telegraphy by using the shortest channel symbol, a dot, for the most common English letter E; while the infrequent letters, Q, X, Z are *represented by* longer sequences of dots and dashes. This idea is carried still further in certain commercial codes where common words and phrases are *represented by* four- or five-letter code groups with a considerable saving in average time.

Shannon (1948, p. 385), emphasis added

Thus the messages of high probability are *represented by* short codes and those of low probability by long codes.

Shannon (1948, p. 402), emphasis added

[The source coding theorem] provides another justification for the definition of entropy rate – it is the expected number of bits per symbol required to *describe* the process.

Cover & Thomas (2006, p. 115), emphasis added

We can design source codes for the most efficient *representation* of the data. [...] The common representation for all kinds of data uses a binary alphabet. Most modern communication systems are digital, and data are reduced to a binary representation for transmission over the common channel.

Cover & Thomas (2006, p. 218), emphasis added

We now discuss the information content of a source by considering how many bits are needed to *describe* the outcome of an experiment.

MacKay (2003, p. 73), emphasis added

Philosophers say that communication theory ignores the content of signals, but communication theorists habitually refer to signal content, using verbs such as ‘represent’ and ‘describe’. As we will see, these locutions play a role in interpreting theorems and justifying mathematical concepts. What’s going on?

Perhaps the tension is only apparent. Perhaps philosophers are aiming for a theory that would attribute content to paradigmatic representations (mental states, neurocognitive processes, linguistic utterances) while communications engineering is concerned with the encoding of digital data. That they both use the term ‘representation’ does not imply that the engineers’ statements are relevant for philosophical theorising. They just happen to use the same term for talking about different kinds of entity in different domains of enquiry.

The problem with this resolution is that the explanatory practices of communication theorists require a relation between signal and signified that looks a lot like representational content (see sections 2 and 3.4). Representationalists take the explanatory role of representation in (for example) cognitive science to warrant realism about content, and design theories to account for that role (Burge, 2010; Shea, 2018). Prima facie, the representationalist should be trying to account for representational explanations in communication theory too.¹ What we actually find in the literature is the view I target here.

1.2 The target view

In this paper I attempt to account for the tension between representation-talk in communication theory and the apparently contradictory assertions of philosophers. I will eventually claim that philosophers haven’t properly distinguished two bearers of representational content in the models of communication theory. To get to that point it would help to first identify a precise claim that the philosophers are making. Clearly they are asserting some kind of negative claim about the connection between communication theory and representational content. The form of that negative claim seems to differ slightly from scholar to scholar: in the quotes above, the theory either “ignores”, “is mute about”, “does not capture”, or “offers no analysis of” representational content. Despite the differences, this kind of claim typically serves a common purpose in justifying the pursuit of a theory of representation that does not involve communication theory. Dretske (1981, p.

¹ Theories of representation in the philosophy of cognitive science typically assert conditions that a cognitive state must meet before it counts as a genuine representation (Burge, 2010; de Souza Filho, 2022; Piccinini, 2020; Shea, 2018). Many such accounts would exclude the communication-theoretic signals discussed in this paper: they are more like simple sensory states than cognitive maps. My contention is that this rejection should take place, if it must, after these signals are recognised as candidates for representational status, just as sensory states are.

65) gives a definition of informational content that is not found in communication theory and bases a theory of semantic content upon it (Dretske, 1981, Chapter 7). Neander (2017, p. 142) gives an analysis of what it means for one event to carry information about another that differs from communication theory and uses it to develop an account of the content of sensory-perceptual representations (Neander, 2017, Chapters 7–9). Dennett (2017, Chapter 6) calls for a theory of semantic information which would be distinct from the tools and concepts of communication theory. Piccinini & Scarantino (2011, sec. 4.2.2) define “non-natural semantic information”, explicitly distinguishing it from Shannon’s formal work.

These authors are trying to justify accounts of representational content that are not constrained by communication theory. On this view, the concepts, methods and results of communication theory play no role in determining in virtue of what a representation has the content it does. Communication theory is irrelevant for philosophical theories of representation. As a result, this paper focuses on the following claim:

THE IRRELEVANCE CLAIM: Communication theory places no constraints on theories of representation.

By ‘theories of representation’ I mean those that are or include a *metasemantic* theory: one that says in virtue of what a representation has the content it does, and perhaps also in virtue of what a putative representation *is* a representation. If THE IRRELEVANCE CLAIM is true, then no theory of representation is ruled out by the practice of communication theory. I think that at least some philosophers would endorse this claim and I think the quotes heading the paper are evidence of this.

I will argue that two popular routes to justifying THE IRRELEVANCE CLAIM are fallacious. In the course of the paper it should become clear that THE IRRELEVANCE CLAIM is false. Communication theory places constraints on metasemantic theories in just the same way other representation-involving sciences do. It thereby rules out any metasemantic theory that fails to attribute the correct representational content to communication-theoretic signals, or which at least gives a deflationary or revisionary account of the representational explanations of communication theorists.² It might appear the philosophers quoted so far are offering a deflationary or revisionary reading of communication theorists’ representation-talk. They aren’t: they are denying such talk exists.

² Although my sympathies lie with the realist who takes communication theorists’ attributions of representational content at face value, I should emphasise that deflationary and revisionary proposals are always live options here. Philosophical work on adjacent disciplines such as neuroscience and computer science has produced legitimate non-semantic readings of representational talk. I cannot argue that communication-theoretic signals must be contentful just because communication theorists describe them as such. Instead I will argue that communication theorists in fact describe such signals as contentful, and that this practice is worthy of philosophical attention. Thanks to an anonymous reviewer for a gentle reminder of this important point.

1.3 Structure of the paper

In section 2 I offer a positive account of representational content. It is a teleosemantic account that defines content in terms of function. I argue that the central model of communication theory is a special case of the basic teleosemantic model, and that teleosemantics therefore agrees with communication theorists on the representational content of signals in their models. My refutation of arguments for The Irrelevance Claim does not depend strictly on the positive view, but I've found it helps to have an account of content on the table.

The majority of the paper focuses on two lines of justification philosophers typically offer in defence of THE IRRELEVANCE CLAIM or claims in its vicinity. I argue that both lines of justification are unsupportable once the formal tools of communication theory, and their contexts of application, are laid out explicitly. In section 3 I examine a warning given by Claude Shannon, the founder of communication theory, that the term 'information' as applied in the theory should be sharply distinguished from the colloquial term 'meaning'. Shannon did indeed make this claim, but it does not imply THE IRRELEVANCE CLAIM, because Shannon was talking about sources, not signals. In section 4, I discuss the fact that certain mathematical tools developed by Shannon can be applied in contexts far removed from signalling systems. Although it is true that mathematical measures like mutual information cannot distinguish the signal-signified relationship from other correlational relationships, this still does not imply THE IRRELEVANCE CLAIM. Communication theory involves models that define signals, and theorems that describe the costs and benefits of transmitting and responding to signals. These theoretical results apply specifically to signalling systems, not just any correlational relationship, and can therefore describe more interesting properties of signals than just quantities of mutual information. Section 5 concludes.

2 The positive view

Before arguing against justifications for THE IRRELEVANCE CLAIM, it can be useful to have a positive account on the table. In this section I describe Millikan's teleosemantics and show that it attributes content to communication-theoretic signals. These attributions agree with the contents communication theorists themselves ascribe to their signals.

2.1 Teleosemantics

Teleosemantics is a rather liberal theory of representation. It attributes representational content to a wide range of artificial and natural systems. Anything that satisfies the model depicted in figure 1 is a representation system, and anything that plays the signal role in that system is a representation. I'll briefly describe

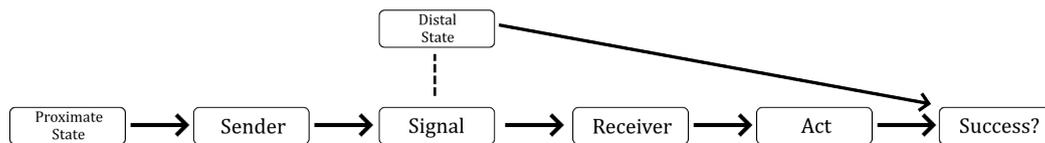


Figure 1: **The basic teleosemantic model.** A **Sender** and **Receiver** have a joint goal they must achieve. This is modelled as the requirement of setting the **Success?** variable to a certain value. The receiver can exercise some control by **Acting**, but a **Distal State** also has causal influence over the target. The sender can produce a **Signal** that bears an exploitable relation to the distal state. When the receiver can condition its behaviour on the signal and achieve greater success than otherwise, teleosemantics claims that explaining this improvement requires appealing to a relation between signal and distal state (dashed line). This is a relation of semantic content: the signal's truth condition is the distal state that must obtain in order for the receiver's act, guided by the signal, to be successful. Finally, how the sender actually produces such a signal is usually by conditioning on one or more upstream **Proximate States**. In special cases, the proximate and distal states are identical (see figures 2 and 3). See also Millikan (2004, fig. 6.3, p. 78).

teleosemantics and how it motivates the claim that representational content is essential to communication theory.

Some artificial and natural devices have the structure and dispositions they do because of causal effects they are supposed to bring about. For artificial devices such as claw hammers, pumps, and communication systems, this link between a device's features (its structure and dispositions) and its proper effects (those it is supposed to bring about) is established by intentional design: a human engineer wanted certain effects to occur, and created the device so as to reliably produce those effects. For natural devices such as claws, hearts, and neurons, the link between features and proper effects is established by natural selection (operating over phylogenetic lineages) and perhaps ontogenic selection (operating within an organism during development). Whichever way the link is established, a device that was designed or selected for producing certain causal effects has a proper function, and its features can be at least partly explained by reference to its function (Millikan, 1984, Chapters 1–2; 1993, Chapter 2).³

When two devices – call them sender and receiver – are designed or selected to assist each other in producing an effect (as in figure 1), an intermediary may enable coordination between them by bearing a relation to a distal state. The distal state causally influences the effect sender and receiver are trying to bring about, so the receiver could be more successful by conditioning its activity on the distal state.

³ As the citations reveal, I'm taking my lead from Millikan's theory of proper functions, on which intentionally produced devices obtain (derived) proper functions from the intentions of their creator. There are a great many issues with this kind of theory that I must ignore for reasons of space. For a recent account, see Garson (2019).

When the receiver uses the intermediary as a proxy and enjoys greater success as a result, its improvement must be explained by reference to a relation between the intermediary and the distal state. Teleosemantics says that when these conditions are met, the intermediary is a representation. A representation has content because it has a distal state to which it is supposed to bear a relation that its receiver can exploit. The corresponding distal state is the representation's truth condition.

2.2 Teleosemantics attributes content to engineered signals

Communication theory concerns devices that have the function of transmitting and responding to signals. In this sense it deals with the very same kinds of system as teleosemantics. The difference between the two theories lies in the questions they attempt to answer: teleosemantics tries to define representations and state their contents, while communication theory tries to devise methods for improving the performance of sender and receiver functions. These questions are typically treated as orthogonal to each other. In section 3.4 I'll argue they are in fact related: improving performance of a signalling system ineliminably involves considerations of signal content. Here I introduce communication theory.

The central engineering model of communication (figure 2) construes the goal of communication as reproducing, at a target location, a symbol string produced at a spatiotemporally distant source. The goal is achieved by **encoding** the source string, which means converting it into a sequence of physical events (typically electrical pulses) that can be transmitted as a signal across a channel. At the far end of the channel the signal is decoded, producing a target string. Communication is deemed successful when the target string is sufficiently similar to the source string. How similar the two strings need to be to count as 'sufficiently similar' will differ depending on the context.

Setting aside noise, the central model is a special case of the basic teleosemantic model. In the teleosemantic model a distal state sits causally upstream of a target that the receiver has causal influence over. The signal has this distal state as its truth condition because the receiver can achieve greater success by conditioning its act on the signal. In the central model, source strings play the role of both distal and proximate states (though more recent models in communication theory distinguish them, e.g. Berger et al. (1996)). Applying the basic teleosemantic model to the central model, the encoder is a sender and the decoder is a receiver. Sender and receiver share a proper function as a consequence of design: to reconstruct the source string. They achieve this goal by means of a signal whose form is determined by a code. Source strings are encoded into signals, and signals are decoded into target strings. Since the encoding defines the relation the signal must bear to the source string in order for the receiver to be successful, teleosemantics agrees with communication theory that the code determines the content for each signal. The truth condition of a signal is the source string it encodes.

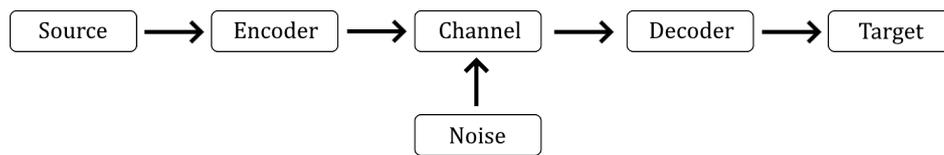


Figure 2: **The central model of communication theory** (see also Shannon, 1948, p. 381 fig. 1). The **Source** is a process that generates strings from a lexicon, within which symbols have a certain probability of appearing. The **Encoder** converts strings of the source lexicon into strings of the code lexicon. The **Channel** is the medium through which code strings are transmitted. During transmission, the code strings are subject to **Noise**, potentially changing their constituent symbols in a non-deterministic way. The **Decoder** attempts to convert code strings back into strings of the original lexicon. Finally, the **Target** is the reconstructed string. The success of communication in this model is measured in terms of the probability of error; specifically, the probability that a symbol in the target string will differ from the corresponding symbol in the source string.

There are a couple of differences between figures 1 and 2 that mean the central model is not quite a special case of the teleosemantic model as depicted. The teleosemantic model does not include noise, but adding a noise variable would not affect the definition of semantic content. The ‘Success?’ variable, and the causal link to it from the relevant distal state, is omitted from the central model; nonetheless, in the central model the receiver’s act together with a distal state determines the joint success of the signalling partnership via an error measure. An example will make this clear.

Consider a system that transmits results of coin tosses $\{H, T\}$. Each time the coin is tossed at the source, the task of the decoder is to produce the appropriate symbol H or T matching the result of the toss. Suppose the sender transmits signals according to the code $H \rightarrow 1, T \rightarrow 0$, and the result of three coin tosses is H, T, H . Then, assuming no noise, when the decoder receives the signal 101 it correctly produces the sequence HTH . Communication is successful because the reproduced string matches the original sequence of results. The signal 101 represents the source sequence HTH , and that is how the decoder successfully reproduces it. Importantly, the claim is not that the content of the signal explains how the receiver produces the string ‘HTH’ at the target. The claim is that the content of the signal explains how the receiver successfully reproduces the source string at the target. In general, teleosemantics claims not that content explains behaviour, but that content explains success. Figure 1 makes that explicit by distinguishing the Act variable from the Success? variable.

Nothing has yet been said about mathematical measures like mutual information or entropy. Neither teleosemantics nor communication theory treats content as identical with, or constructed from, correlational relationships. Rather, it is the encoding scheme that determines the content of a signal. The reason the encod-

ing scheme determines content is that it is the encoding scheme that determines which signals will prompt which receiver acts, hence which relations between signal and source string will enable the receiver to enjoy increased success in its goal of reproducing the source string. We tend to think of encoding as the translation of a proximate state into a signal. In the central model, encoding also determines relations borne between signal and distal state. That's because the distal state just is the proximate state in that model.

So far, I have briefly surveyed positive reasons for thinking communication theory involves considerations of representational content (more are given in section 3.4). This conflicts with The Irrelevance Claim. Instead of fleshing out the positive view, the rest of the paper aims to undercut the scepticism. The next two sections describe and refute two popular sceptical arguments about the relationship between communication theory and representational content.

3 First sceptical argument: SHANNON'S WARNING

The first route to justifying THE IRRELEVANCE CLAIM begins with a warning given by Claude Shannon, the founder of communication theory, that his theory had nothing to do with meaning. In this section I demonstrate that Shannon's warning pertains to sources, not signals. When Shannon did turn to signals, he called them representations and explicitly referred to them as contentful. Contemporary communication theorists endorse this pretheoretic attribution of content, and the standard interpretation of one of the theory's fundamental theorems requires it.

3.1 In the central model, efficient signalling depends on source probabilities

The heavy cryptographic and communicative demands of the Second World War led mathematicians and engineers, spearheaded by Claude Shannon, to develop the discipline known today as communication theory. Published soon after the war's conclusion, the insights of Shannon's foundational text (1948) are predicated on the central model (figure 2). Recall that the goal of communication in this model is to reproduce a source string at a target location. Whether or not a source string is meaningful is irrelevant to the task at hand. Suppose for example the lexicon from which the source string is constructed is the Latin script $\{A, B, C \dots\}$ plus a full stop and a space. The code lexicon might be the binary symbols $\{0, 1\}$ that are instantiated by electrical on/off pulses. An encoding scheme converts each string of Latin symbols into a sequence of 0s and 1s which is then transmitted as a signal across a wire. The decoder has a duplicate set of Latin symbols from which it must pick out the right symbols in the right order; the signal enables it to perform this task successfully.

As an engineering science, communication theory is concerned with reconstructing symbol strings efficiently, which means encoding and transmitting

signals for as little cost as possible. Electrical wires require power and time to operate. Communication theory can be thought of as a collection of tools and methods enabling an optimal trade-off between signalling effort and the benefits of accurate string reconstruction. In the central model the most efficient coding schemes are those that use short sequences of 0s and 1s to represent highly probable source strings. That is because in that simplified setting minimising signalling effort means minimising the number of code symbols transmitted, on average. Pairing probable source strings with short code strings – short signals – is the most efficient procedure. Therefore, in order to devise a good code, you need to know the probabilities of each source string being produced. Crucially, that is *all* you need to know. Whether or not source strings also carry natural language meaning is irrelevant to the problem of designing a code. Shannon stated this clearly, as the next subsection details.

3.2 SHANNON'S WARNING

In the introduction to the first of his foundational papers Shannon writes:

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem.

Shannon (1948, p. 379), emphasis original

Clearly “message” in this context refers to a source string. Shannon warns that the semantic properties of lexical elements do not affect the process of transmitting and reconstructing them. To see why this is true, note that strings of Latin symbols need not form English-language words, nor words of any other language. The problem of reconstructing those strings has a distinct mathematical sense, regardless of the strings’ natural language implications.

In 1949 Shannon’s papers were released in a single volume with prefatory remarks by Warren Weaver (Shannon & Weaver, 1949). One of Weaver’s comments expands on Shannon’s earlier technical statement:

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. It is this, undoubtedly, that Shannon means when he says that “the semantic aspects of communication are irrelevant to the engineering aspects.”

Shannon & Weaver (1949, p. 8)

Weaver misquotes Shannon (“*the* semantic aspects” instead of “*these* semantic aspects”; “the engineering *aspects*” instead of “the engineering *problem*”). In context the mistake is insignificant, because the preceding sentences demonstrate that Weaver interprets the point accurately. By ‘information’ he is referring specifically to a property of source strings called **surprisal**. Surprisal is defined as $\log \frac{1}{p(x)}$ for a message x that is produced at the source with probability $p(x)$. (Surprisal is a useful measure of a source string because it is equal to the number of symbols in the corresponding code string given an optimal encoding scheme.) It is clear that two messages – two source strings – can have the same surprisal, with one being a meaningful sentence of a natural language and the other being nonsense. One need only define a source that produces a meaningful sentence and a nonsensical sentence with the same probability.

In context the misquote is unproblematic, but out of context Weaver can be read as stating what I will eventually deny: that the semantic properties of both the source string *and the code string* are irrelevant to the well-functioning of engineered communication systems. To dispel any doubt, I endorse Shannon’s original claim. I take it to be as follows:

SHANNON’S WARNING: In the central model, once the statistical properties of source strings have been taken into account, the semantic properties of source strings are irrelevant to the engineering problem of communication.

The meanings of source strings are not represented in the mathematics of communication theory. SHANNON’S WARNING tells us that finding an efficient solution to the fundamental problem of communication requires knowing only the statistical properties of source strings, not their meanings.

Philosophers took heed of SHANNON’S WARNING. Over the years, however, it mutated into a different claim.

3.3 Philosophers’ interpretations of the warning

Soon after Shannon’s initial publications, Bar-Hillel & Carnap (1953) set the standard for philosophical interpretation of communication theory:

The Mathematical Theory of Communication, often referred to also as Theory (of Transmission) of Information, as practised nowadays, is not interested in the content of the symbols whose information it measures. The measures, as defined, for instance, by Shannon, have nothing to do with what these symbols symbolise, but only with the frequency of their occurrence.

Bar-Hillel & Carnap (1953, p. 147)

Like Weaver, Bar-Hillel and Carnap likely understood SHANNON'S WARNING correctly; like Weaver, the words they used could be misconstrued. By referring only to 'symbols' they risked conflating source symbols and code symbols. Nonetheless, the job of philosophers, as Bar-Hillel and Carnap saw it, was to provide a theory of *semantic information* that would capture the aspect Shannon ignored. They explicitly distinguish two kinds of theory, implying a distinction between two entities or concepts.

Bar-Hillel and Carnap's exposition had significant influence. Dretske (1981, p. 241, n. 1) cited them as the best-known sceptics about the relevance of communication theory for philosophical questions about content. Dretske also offered an interpretation of SHANNON'S WARNING:

Communication theory does not tell us what information is. It ignores questions having to do with the *content* of signals, what *specific information* they carry, in order to describe *how much* information they carry. In this sense Shannon is surely right: the semantic aspects are irrelevant to the engineering problems.

Dretske (1981, p. 41), emphasis original

Two things are worth noticing. First, Dretske is talking about the content of *signals*, whereas SHANNON'S WARNING concerns the semantic properties of source strings. Second, Dretske repeats Weaver's misquote of Shannon: "*the* semantic aspects" instead of "*these* semantic aspects" (strictly speaking, Dretske does not use quotation marks – but earlier on the same page he repeats the misquote along with an end-note reference to Shannon's original statement). Influenced by Dretske, Dennett (1983) repeated Bar-Hillel and Carnap's call for a distinction between mathematical and semantic information:

A more or less standard way of introducing the still imperfectly understood distinction between these two concepts of information is to say that Shannon-Weaver theory measures the *capacity* of information-transmission and information-storage vehicles, but is mute about the *contents* of those channels and vehicles, which will be the topic of the still-to-be-formulated theory of semantic information.

Dennett (1983, p. 344), emphasis original

Dennett speaks of "channels and vehicles", which would presumably include signals. Like Dretske, the version of the warning operative here is different from the original statement. Dennett (2017, Chapter 6) is still pursuing this line, and it is nowadays standard to distinguish between two senses of the term 'information' in scientific applications. The Stanford encyclopedia entry 'Biological Information' is organised around the distinction, using the labels "Shannon's concept of information" and "Teleosemantic and other richer concepts" (Godfrey-Smith & Sterelny,

2016). Piccinini & Scarantino (2011, p. 21) say “Shannon information does not capture, nor is it intended to capture, the semantic content, or meaning, of signals,” again focusing on signals rather than source strings. It is accepted practice to refer to Shannon’s formal tools as unrelated to semantic content without further argument: “I will interpret ‘information’ as ‘semantic information’ (i.e. semantic content), not as Shannon information” (Artiga, 2021, p. 12 n. 14); see also Cao (2020, p. 6) and Kolchinsky & Wolpert (2018, p. 1).

The distinction between two concepts of information reinforces The Irrelevance Claim. It motivates the view that philosophers are interested in one thing (semantic information, usually treated as synonymous with or related to semantic content) while communication theorists are interested in a distinct and unrelated thing (Shannon information). A charitable way to understand this view is that it distinguishes semantic content from correlational measures like mutual information. The philosophers just cited could be read as claiming that signals can bear mutual information without possessing semantic content. I discuss such claims in section 4; my point here is that these writers cannot appeal to Shannon and Weaver to justify their position. Weaver’s quote above uses ‘information’ as synonymous with surprisal, not mutual information; both he and Shannon were referring to a property of source strings, not a correlation between signals and signifieds. To my knowledge, no explicit argument has been offered that moves from SHANNON’S WARNING to something resembling THE IRRELEVANCE CLAIM. Most likely there is no valid argument of this kind, as it would require moving from the true claim that source strings need not have semantic content to the false claim that signals need not have semantic content.

Perhaps the strongest argument for my position is that The Irrelevance Claim is false by the lights of communication theorists themselves. Signals in the central model have semantic content, and their content is directly relevant to the engineering problem of communication, as we shall now see.

3.4 The truth condition of a signal in the central model is a source string

Contemporary discussion of the relevance of communication theory for semantic content focuses on signals. But nothing about signals can be concluded directly from SHANNON’S WARNING, which concerns source strings only. The mathematical tools of communication theory are indeed blind to the meaning of source strings, but interpretations of the theory require that signals are contentful: the truth condition of a signal is a source string. (For simplicity I assume a one-to-one mapping between source strings and code strings; my substantive points are not affected by loosening this constraint.) We’ve already seen quotes from communication theorists attributing content to signals, and I’ve offered teleosemantics as a theory that takes this usage seriously. Here I will give more context to those quotes in order to strengthen the plausibility of my interpretation.

Describing Morse code, Shannon (1948, p. 385) says that letters are “represented by” sequences of dots and dashes. In their widely cited textbook Cover & Thomas (2006, p. 105) say the same. Shannon further uses the locution “represented by” in this context on pages 402 and 405; Cover & Thomas use this and related terms (including calling encodings “representations”) on pages 5-6, 130, 134, 218-9, 221 and 301. Obviously the authors cannot be interpreted as taking a stance on philosophical theories of content. Nevertheless, their usage is evidence of a pretheoretic notion of representation at work, a relation of signification linking signals with source strings.

Furthermore, standard interpretations of the theorems of communication theory require that signals signify source strings. Consider the **first theorem of communication theory**, often called the source coding theorem (Cover & Thomas, 2006, Chapter 5; MacKay, 2003, Chapter 4). The theorem gives a lower bound on the average number of binary code symbols required to encode source strings of a specified length. The fewest code symbols required is a function of the **entropy** of the source, which itself is a function of the probabilities of source symbols. The theorem clearly embodies a notion of signification or reference: both the question that prompted the theorem (how many code symbols are required?) and the result it offers (the entropy of the source) assume that the code symbols are being used to record the symbols of the source string. Cover & Thomas make this clear:

This theorem provides another justification for the definition of entropy rate – it is the expected number of bits [i.e. code symbols] per [source] symbol required to describe the [source] process.

Cover & Thomas (2006, p. 115)

The extent to which Cover & Thomas’s use of terms like “describe” and “represent” corresponds with philosophers’ notions of semantic content has not to my knowledge been asked. Their usage is evidence that some notion of signification is required to make sense of the theorem. A similar sentiment is found in MacKay’s textbook, wherein he models a scientist’s experimental setup as a source and the different possible outcomes of the experiment as different source strings:

We now discuss the information content [entropy] of a source by considering how many bits [code symbols] are needed to describe the outcome of an experiment.

MacKay (2003, p. 73)

Again, the sense of the entropy measure is intimately bound up with representation, in this case in describing the outcome of the experiment. Similarly, in a discussion of a special case of the theorem, Shannon (1948, p. 397) speaks of “the number of bits [code symbols] required to specify the sequence” of symbols in a source string.

Overall, communication theorists from Shannon to contemporary textbook authors make use of a notion of semantic content, as indicated by their use of words like ‘represent’, ‘describe’ and ‘specify’. There is an intuitive sense to this usage. Signals in the central model must bear some exploitable relation to source strings, because the decoder is using the signals to reconstruct the original strings. The specific mapping from signal to source – the ‘semantics’ of the signalling system – is determined by the encoding scheme.

3.5 A sceptical riposte: symbol manipulation does not bestow content

Encoding, the sceptic will notice, is the transformation of symbols from one lexicon into another. Since I am claiming it is the encoding scheme that confers content, my position appears to entail that any process by which symbols of one lexicon are converted into symbols of another confers content. That is a problem: it is implausible that manipulating symbols from lexicon L_1 into lexicon L_2 bestows the symbols of L_2 with the content ‘symbol such-and-such from L_1 ’. Symbol manipulation is a matter of syntax, not semantics. If that is all that is happening in an encoding scheme, then it is implausible that central model signals really do have the contents I ascribe.

To respond, I accept the premise of the objection. Converting symbols from one lexicon to another does not bestow them with content. The question is whether symbol manipulation is all there is to encoding. While the term ‘encoding’ might be used in different ways in different branches of science and philosophy, including in ways that imply only symbol manipulation, I contend that its use in communication theory implies something stronger. Encoded strings are produced as part of a sender-receiver system, in order to be decoded during performance of a joint function. That makes a difference because it ensures the system fits the teleosemantic template. Shuffling symbols does not bestow content but joint design of sender and receiver does.

Furthermore, although I have focused on symbols in the exposition so far, it turns out that source ‘symbols’ need not be symbols at all. They could be dance steps, military manoeuvres, restaurant locations; any outcome from an event space over which a probability distribution can be defined. It also turns out that the actions of the receiver need not be exact duplicates of the outcomes at the source; they need only be actions that, combined with source outcomes, yield a cost (via a cost function) for the system as a whole. If this sounds like the sender-receiver framework associated with Skyrms (2010) and Lewis (1969) that’s because it is formally equivalent to it (Martínez, 2019b). Sender-receiver games (figure 3) are also a special case of the basic teleosemantic model, and the central model just is a sender-receiver game. Source and target – state and act – need not be symbols. They are just commonly described as such because that is part of the typical use

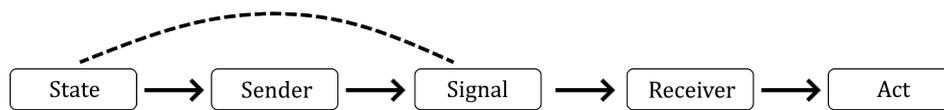


Figure 3: **A signalling game.** In cooperative settings, **Sender** and **Receiver** must collaborate to achieve a shared payoff. The payoff depends on the receiver's **Act** and the **State** observed by the sender. Since only the sender has access to the state, it must guide the receiver with a **Signal**. Here, the state plays both roles that teleosemantics distinguishes as proximate and distal states (see figure 1): it is a proximate state because the sender conditions its choice of signal on it, and it is a distal state because the value of the target (the payoff) depends on it. Teleosemantics says that the truth condition of a signal is the corresponding state (dashed line). Signalling games include payoff matrices that yield reward values from combinations of states and acts; these could instead be represented by a downstream Success? variable as in figure 1. See also Martínez (2019b) for the close formal links between communication theory and signalling games.

case of communication theory. The mathematics does not demand that signals be about symbols from a lexicon. They can be about anything at all.

To summarise the entire section, arguments denying the relevance of communication theory for theories of content based on SHANNON'S WARNING do not hold water. This route to THE IRRELEVANCE CLAIM is blocked. A stronger argument adverts to the breadth of application of another of Shannon's mathematical measures: mutual information. It is to this point I now turn.

4 Second sceptical argument: Agnostic information

The second route to THE IRRELEVANCE CLAIM begins with the idea that mathematical measures defined within communication theory cannot distinguish between representations and non-representations. In other words, information is agnostic to representational status. In this section I argue that although this is true of certain mathematical functions like mutual information, communication theory as a whole does distinguish between signals and non-signals – it must do in order for its theorems to have sense. The theory is fundamentally about the costs and benefits of representation and how to trade them off judiciously. Focusing on mutual information obscures this wider theoretical perspective.

4.1 How scientists use information theory

Soon after Shannon's original text, scientists began to notice that his mathematical tools were of use beyond the context of communications engineering. Perhaps the

most well-known informational measure is **mutual information**, typically interpreted as the strength of correlation between two variables. Mutual information has been employed in a diverse range of sciences, including:

- Behavioural ecology, to measure the correlation between the honeybee waggle dance and the location of food sources (Haldane & Spurway, 1954)
- Cosmology, to measure the correlation between galaxies' internal morphology and their local environments (Pandey & Sarkar, 2017)
- Evolutionary biology, to show that the correlation between an environmental cue and a fitness-relevant state of affairs is an upper bound on the increased growth rate of an organism conditioning its behaviour on the cue (Donaldson-Matasci et al., 2010)⁴
- Linguistics, to measure the co-occurrence of words in a corpus (Hunston, 2002, Chapter 4)
- Molecular biology, to measure the correlation between inputs and outputs of a quorum-sensing bacterium (Mehta et al., 2009)
- Neuroscience, to measure the correlation between neural firings and environmental states (Rathkopf, 2017a and references therein)

This multifaceted role for mutual information accompanied the emergence of the term **information theory** to describe Shannon's mathematical tools and their more general application across the sciences. Today, information theory comprises a set of concepts and measures common to many mathematical and scientific disciplines (Cover & Thomas, 2006, p. 2, fig. 1.1). It has become customary for philosophers to use the term 'information theory' to cover all applications of informational measures and concepts, including in communicative settings. The claims I make below depend on there being an important distinction between information theory in general and the specific formal tools brought to bear on communicative partnerships like the central model (recall figure 2). As such, I'll use the term 'information theory' in the customary very broad sense and 'communication theory'

⁴ How can a correlation be an upper bound on increased growth rate? An anonymous reviewer pointed out that the units of correlation and growth rate appear to be different, so one cannot be an upper bound on the other. Surprisingly, Donaldson-Matasci et al. (2010, pp. 224–226) show that when increased growth rate is measured as the difference in expected log fitness, its units become commensurate with informational units (e.g. bits), and (given certain further assumptions) that the mutual information between a cue and a fitness-relevant environmental state does indeed become an upper bound on increased growth rate. There's a lot to unpack here, and philosophical work on the propriety of their model is lacking. I'm just using their result as an example of the diversity of applications of mutual information across the sciences.

to mean the set of formal tools, concepts and methods developed by engineers and scientists to investigate communication from a mathematical point of view.⁵

The breadth of application of mutual information is at the heart of a second source of scepticism about the relevance of communication theory for theories of content. Sceptics seem to move from a premise about mutual information to a conclusion about information theory as a whole – which is then seen as encompassing communication theory. In the next two subsections I argue that the premise is true but the conclusion is false; no valid argument can link them.

4.2 Mutual information cannot distinguish signals and cues

Originating in behavioural ecology, the signal/cue distinction highlights the fact that some informational vehicles have the function to provide the information they do, whereas some are ‘accidentally’ informational, used opportunistically by their receivers (Maynard Smith & Harper, 2003, sec. 1.2). Vehicles selected to serve a communicative role are called signals, while vehicles that fortuitously provide information are called cues. (In philosophy the term ‘natural sign’ is sometimes used; I am here using ‘cue’ to cover all cases described by that term.) The waggle dance is a signal because it evolved in the honeybee lineage to serve as an informational vehicle that enables workers to enjoy greater success at foraging or nest-finding (Gould, 1975; Riley et al., 2005). In contrast, bees’ use of the position of the sun in the sky to navigate is a cue, because the sun’s location is not a result of a process of selection that jointly produced both it and the bees’ navigational behaviour.

Mutual information quantifies the strength of a correlation no matter whether its vehicles are signals or cues. The sheer variety of scientific contexts employing mutual information emphasises this point. While the correlation between the waggle dance and food locations is due to the fact that the waggle dance is a signal, the correlation between galaxies’ morphology and their local environment clearly is not. The vehicle in the evolutionary model of Donaldson-Matasci et al. (2010) is definitionally a cue. The co-occurrence of words in a corpus is not a signal (though the words themselves are representations, or at least combine to produce representations). Without further detail, it is not clear whether the output of a quorum-sensing bacterium counts as a signal of its input; nonetheless, mutual information between the two can be measured. Neural firings are sometimes claimed to be representations, but simply measuring the correlation between them and environmental states is not sufficient to establish this (Rathkopf, 2017a).

From the mere fact that two things bear a correlational relationship, no conclusion can be drawn about whether one is a signal of the other. I fully agree with this premise and suggest encapsulating it as follows:

⁵ As an anonymous reviewer helpfully pointed out, my argument is therefore not restricted to scenarios communication theorists themselves happen to be interested in, but any application of informational concepts to functional sender-receiver partnerships, as in figure 2.

MUTUAL INFORMATION IS AGNOSTIC: Mutual information cannot distinguish signals and cues.

The path to scepticism I want to explore is the move from MUTUAL INFORMATION IS AGNOSTIC to the following claim:

COMMUNICATION THEORY IS AGNOSTIC: Communication theory cannot distinguish signals and cues.

The latter claim would lend strong support to THE IRRELEVANCE CLAIM. If communication theory cannot distinguish signals and cues, there is little hope of it saying anything about content.⁶

I am going to argue that COMMUNICATION THEORY IS AGNOSTIC is false. Communication theory can and does distinguish signals and cues. If space permitted I would attempt to reconstruct the arguments of philosophers who appear to move from the true premise to the false conclusion. Space does not permit, so I will instead note that several writers appear to trade on an ambiguity in the term ‘Shannon information’, treating it as both a synonym for mutual information and a catch-all that exhausts the conceptual repertoire of communication theory (Godfrey-Smith & Sterelny, 2016, p. 1 and §2; Dennett, 2017, p. 106; Owren et al., 2010, p. 759; Shea, 2018, p. 12, n. 11 and p. 78, n. 5). Martínez (2019a, p. 1216) has similarly argued that philosophers routinely treat communication theory as providing no more conceptual tools than measures of correlation.

4.3 COMMUNICATION THEORY IS AGNOSTIC is false

Whereas information theory is a collection of mathematical tools with wide application across the sciences, communication theory is an engineering discipline with the specific goal of designing efficient signalling techniques. The vehicles transmitted in the central model are signals and the main theorems of communication theory apply to signals. Consider, on the one hand, the **second theorem** and **third theorem** of communication theory, which require that the vehicles they address be signals, and on the other hand **Kelly’s theorem**, which requires only that the vehicle be a cue. We shall introduce them in turn.

The second theorem of communication theory (sometimes called the noisy channel theorem or the channel coding theorem) determines how accurately a receiver is able to reconstruct a source string given a signal that has been corrupted by noise. Better encodings combat noise by building redundancy into the signal,

⁶ Here I assume cues do not have semantic content. This assumption is popular but has recently been challenged by Isaac (2018) and Skyrms (2010, Chapter 3) who define content in terms of correlational relationships. They carefully deconstruct the formal components of correlational measures to determine relational properties of individual signs. Space precludes discussion of this ingenious alternative. My theoretical commitments lead me to assign content to signals only; the reader may have different inclinations.

enabling the receiver to more accurately reconstruct the source. The theorem answers a question about receiver performance by attending to the sender's design of the vehicle: different encoding schemes would yield different performance levels. The vehicle mentioned in this theorem is definitionally a signal; the theorem does not apply to cues.

The third theorem of communication theory (also known as the rate-distortion theorem) addresses a similar question, this time with the added benefit that the receiver need not achieve perfect performance. Suppose for example that the receiver only needs to correctly reconstruct four out of every five outcomes produced by the source. The third theorem states that it is possible to determine the minimum transmission rate that the sender must ensure in order for the receiver to perform at the specified level. The theorem assumes that the transmission rate is tunable by the sender's choice of encoding scheme. By invoking a vehicle whose form can be adapted to performance specifications, both the second and third theorems employ a concept of signal rather than cue.

Kelly's theorem, by contrast, concerns the performance of a receiver conditioning its behaviour on a vehicle that bears a fixed level of mutual information with a success-relevant distal state (Kelly, 1956). The theorem states that mutual information is an upper bound on the performance improvement a receiver can enjoy by using the vehicle rather than not using it. Because the emphasis is on receiver behaviour, and the mutual information is held fixed, the vehicle is conceptualised as a cue. The theorem has been applied in evolutionary biology (Donaldson-Matasci et al., 2010) where again the vehicle in question is treated as an environmental cue rather than a signal.

Of course, nothing prevents us applying Kelly's theorem to signals too. My claim is not that theorems about cues do not apply to signals, but that theorems about signals do not apply to cues. There is an asymmetry in the definition of signals and cues, and the two kinds of theorem are asymmetric in a way that reflects that. Any vehicle can be treated as a cue simply by failing to specify whether or not it was designed for communicative use. That is what Kelly's theorem does, in the guise of keeping mutual information fixed and asking how the receiver can make use of it. (In fact Kelly's prose implies that the vehicle in question is a signal; closer inspection reveals the theorem does not require it to be one.) The second and third theorems by contrast require that their vehicles be signals, because they ensure a level of functional performance that is only available when the sender tunes the vehicle's production to the features of the channel. By definition, the second and third theorems cannot be applied to cues.

Another way to understand the difference here is to note that signals enter into certain causal relationships that cues need not. That a signal forms a causal bridge between sender and receiver is what enables the sender to implement an encoding scheme, intervening on the causal order in such a way as to improve functional performance. Since measuring mutual information alone implies nothing about causality, communication theory must be employing more formal tools than just

mutual information when it is describing signals. Calcott et al. (2020) make a similar point with reference to Skyrms's (2010) definition of information in signals. Skyrms describes an informational property of signals that is related to mutual information and, like mutual information, has no implications about the causal relationships a signal is entering into. Calcott and colleagues argue that this informational property therefore misses an important fact about signals; namely that they do not just carry information about their effects but are difference-makers for their effects. By adding formal apparatus derived from Woodward's (2003) interventionist account of causal explanation, Calcott and colleagues augment Skyrms's account and enable signals to be distinguished from mere cues. The point that's relevant for our purposes is that the central model, like all sender-receiver models, is by definition causal. There must be a causal link between sender and receiver in order for there to be a signalling partnership in the first place. The second and third theorems require a causal assumption that Skyrms's informational property, like mutual information, ignores. Rathkopf (2017a) points out similar issues with mutual information and causal assumptions in the case of neurobiology.

In sum, communication theory distinguishes signals from cues both by providing the means to define signals and by employing theorems that require the vehicles in question to be signals. The fact that communication theory also contains theorems like Kelly's whose vehicles need only be cues serves to sharpen the point. COMMUNICATION THEORY IS AGNOSTIC is false, and this route to establishing THE IRRELEVANCE CLAIM is blocked.

4.4 A sceptical riposte: channel capacity is agnostic about content

Alarm bells are ringing in sceptical ears: I just talked about the second theorem, and how the requirement of a causal link between sender and receiver confirms its vehicles are signals. But the sceptic knows that the second theorem deals with the **capacity** of the channel. The capacity is defined only in terms of the signal before noise and the signal after noise; it is oblivious to the source and target strings – oblivious to what the signal is actually about. If the second theorem and the channel capacity it defines are agnostic to the content of signals, is not the sceptic justified in asserting that the definition of semantic content must be found elsewhere than communication theory? Dennett (1983, p. 344) seems to be making this point when he distinguishes between the “*capacity* of information-transmission and information-storage vehicles” and their contents, stating that Shannon's theory deals only with capacity. More recently he reasserts the claim:

Shannon devised a way of *measuring* information, independently of what the information was *about*, rather like measuring *volume* of liquid, independently of which liquid was being measured. (Imagine someone bragging about owning lots of quarts and gallons and not

having an answer when asked, “Quarts of what – paint, wine, milk, gasoline?”)

Dennett (2017, p. 106), emphasis original

To respond, I begin by accepting one of the premises of the sceptical riposte. It is true that channel capacity can be measured without specifying the contents of signals transmitted through the channel. Channels, characterised by the conditional distribution of signal-after-noise given signal-before-noise, are general-purpose in that the outcomes of any source can be encoded by the code symbols. Storage devices like hard drives, and transmission media like fibre-optic cables, can be assigned measures of capacity without regard to what they are storing or transmitting. Any channel can in principle be used to communicate anything. The honeybee waggle dance could be used to communicate military instructions if the field commander sending the message had sufficiently fine-grained control over the placement of food sources in the bees’ locale. Measuring the capacity of the dance cannot tell you what, on a given occasion, is being communicated by it.

However, the general-purpose nature of channels does not entail that communication theory cannot attribute content to signals. Signals in the central model (for example) do have content as soon as a source and an encoding scheme are specified: the contents of signals are the different outcomes of the source. If an army managed to employ the waggle dance to transmit military instructions, each dance would gain a specific instruction as one of its semantic contents in accordance with the code devised ahead of time by the human communicators. Furthermore, measurements of transmission rate would reveal how fine-grained the field commander’s control of food source placement would have to be in order to transmit distinguishable instructions. Those measurements could contribute to a determination of whether or not this system was worth employing as a means of communication.

The upshot, as with several other arguments discussed in this paper, is that a true premise has been used to derive a false conclusion. That a channel can be used to transmit anything does not mean communication theory remains forever agnostic about what is actually being transmitted in a given circumstance. In terms of Dennett’s analogy, the theory both measures the volume of a liquid and tells you what liquid is being measured. Dennett really does appear to be saying that Shannon’s formal work encompasses only measures like mutual information and channel capacity. If he means to say this, he is wrong.

To summarise the entire section, true premises about the broad application of mutual information in science and the general-purpose nature of channels provide no reason to be sceptical about the relevance of communication theory for theories of content.

5 Conclusion

THE IRRELEVANCE CLAIM says that communication theory places no constraints on theories of representational content. Many philosophers try to establish this claim, or others in its vicinity, by following two routes that each begin with a true claim. SHANNON'S WARNING and MUTUAL INFORMATION IS AGNOSTIC are both true. However, neither of them suffices to establish the sceptical conclusion. Perhaps there are other routes to THE IRRELEVANCE CLAIM. Given the explanatory role of signal-signified relations in communication theory, this seems doubtful. We should treat the locutions of engineers like we treat the locutions of cognitive scientists: apt for justification, deflation, or revision. In any case we ought not ignore them.

At the beginning of the paper I argued that the central model of communication theory conforms to the teleosemantic template, and that therefore communication-theoretic encodings are among the content-determining relations teleosemantics describes. This consilience suggests a change to the standard view that treats formal approaches to communication as orthogonal to philosophical issues of semantic content. If encoding schemes imbue signals with semantic content, it will pay philosophers to investigate the different kinds of encoding scheme that can exist, and the breadth of application of communication-theoretic models across the functional sciences. Perhaps light can be shed on debates in disciplines such as neuroscience by applying the models of communication theory to neural signalling networks. If neuroscientists could get mileage from this approach, the connection between communication theory and teleosemantics would augment existing philosophical work in this area (Cao, 2012, 2014; Rathkopf, 2017a, 2017b). Overall, the idea that formal tools are conceptually and technically distinct from the philosophical question of semantic content is becoming untenable.

Acknowledgments

Thanks to Brant Pridmore and Jeremy Strasser for comments on archaic drafts. The present version was developed for the Representation in Cognitive Science workshop at the University of Barcelona in September 2022, and I am grateful to the organisers and participants for the chance to present and for the helpful feedback. Subsequently, Johan Heemskerk, Manolo Martínez, Karl Bergman, Sergio Balari and Oriol Roca-Martín provided invaluable discussion on a manuscript that looked recognisably like this one but was much improved by their input. Finally, two anonymous reviewers and the editor of the present journal contributed positively to the text as you see it today.

This work was supported by Juan de la Cierva grant FJC2020-044240-I and María de Maeztu grant CEX2021-001169-M funded by MCIN/AEI/10.13039/501100011033.

References

- Artiga, M. (2021). Signals are minimal causes. *Synthese*, 198, 8581–8599. <https://doi.org/10.1007/s11229-020-02589-0>
- Bar-Hillel, Y., & Carnap, R. (1953). Semantic Information. *The British Journal for the Philosophy of Science*, 4(14), 147–157. <http://www.jstor.org/stable/685989>
- Berger, T., Zhang, Z., & Viswanathan, H. (1996). The CEO problem [multiterminal source coding]. *IEEE Transactions on Information Theory*, 42(3), 887–902. <https://doi.org/10.1109/18.490552>
- Burge, T. (2010). *Origins of Objectivity*. Oxford University Press.

Mann, S. F. (2023). The relevance of communication theory for theories of representation. *Philosophy and the Mind Sciences*, 4, 32. <https://doi.org/10.33735/phimisci.2023.10992>



©The author(s). <https://philosophymindscience.org> ISSN: 2699-0369

- Calcott, B., Pocheville, A., & Griffiths, P. E. (2020). Signals That Make a Difference. *The British Journal for the Philosophy of Science*, 71(1), 233–258. <https://doi.org/10.1093/bjps/axx022>
- Cao, R. (2012). A teleosemantic approach to information in the brain. *Biology & Philosophy*, 27(1), 49–71. <https://doi.org/10.1007/s10539-011-9292-0>
- Cao, R. (2014). Signaling in the Brain: In Search of Functional Units. *Philosophy of Science*, 81(5), 891–901. <https://doi.org/10.1086/677688>
- Cao, R. (2020). New Labels for Old Ideas: Predictive Processing and the Interpretation of Neural Signals. *Review of Philosophy and Psychology*, 11, 517–546. <https://doi.org/10.1007/s13164-020-00481-x>
- Cover, T. M., & Thomas, J. A. (2006). *Elements of Information Theory* (2nd ed.). John Wiley & Sons.
- de Souza Filho, S. F. (2022). A Dual Proposal Of Minimal Conditions For Intentionality. *Synthese*, 200(2), 115. <https://doi.org/10.1007/s11229-022-03483-7>
- Dennett, D. C. (1983). Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. *Behavioral and Brain Sciences*, 6(3), 343–355. <https://doi.org/10.1017/S0140525X00016393>
- Dennett, D. C. (2017). *From Bacteria to Bach and Back: The Evolution of Minds*. Penguin UK.
- Donaldson-Matasci, M. C., Bergstrom, C. T., & Lachmann, M. (2010). The fitness value of information. *Oikos*, 119(2), 219–230. <https://doi.org/10.1111/j.1600-0706.2009.17781.x>
- Dretske, F. I. (1981). *Knowledge and the flow of information*. MIT Press.
- Garson, J. (2019). *What Biological Functions Are and Why They Matter*. Cambridge University Press.
- Godfrey-Smith, P., & Sterelny, K. (2016). Biological Information. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2016, pp. 1–37). The Metaphysics Research Lab. <http://plato.stanford.edu/archives/sum2016/entries/information-biological/>
- Gould, J. L. (1975). Honey bee recruitment: The dance-language controversy. *Science*, 189(4204), 685–693. <https://doi.org/10.1126/science.1154023>
- Haldane, J. B. S., & Spurway, H. (1954). A statistical analysis of communication in "Apis mellifera" and a comparison with communication in other animals. *Insectes Sociaux*, 1(3), 247–283. <https://doi.org/10.1007/BF02222949>
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524773>
- Isaac, A. M. C. (2018). The Semantics Latent in Shannon Information. *The British Journal for the Philosophy of Science*, 70(1), 103–125. <https://doi.org/10.1093/bjps/axx029>
- Kelly, J. L. (1956). A New Interpretation of Information Rate. *Bell System Technical Journal*, 35(4), 917–926. <https://doi.org/10.1002/j.1538-7305.1956.tb03809.x>
- Kolchinsky, A., & Wolpert, D. H. (2018). Semantic information, autonomous agency and non-equilibrium statistical physics. *Interface Focus*, 8(6), 20180041. <https://doi.org/10.1098/rsfs.2018.0041>
- Lewis, D. (1969). *Convention: A philosophical study*. Blackwell.
- Lombardi, O., Holik, F., & Vanni, L. (2015). What is Shannon information? *Synthese*, 193(7), 1983–2012. <https://doi.org/10.1007/s11229-015-0824-z>
- MacKay, D. J. (2003). *Information theory, inference and learning algorithms*. Cambridge University Press. <http://www.inference.org.uk/mackay/itila/book.html>
- Martínez, M. (2019a). Representations are Rate-Distortion Sweet Spots. *Philosophy of Science*, 86(5), 1214–1226. <https://doi.org/10.1086/705493>
- Martínez, M. (2019b). Deception as cooperation. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 77, 101184. <https://doi.org/10.1016/j.shpsc.2019.101184>
- Maynard Smith, J., & Harper, D. (2003). *Animal Signals*. OUP Oxford.
- Mehta, P., Goyal, S., Long, T., Bassler, B. L., & Wingreen, N. S. (2009). Information processing and signal integration in bacterial quorum sensing. *Molecular Systems Biology*, 5, 325. <https://doi.org/10.1038/msb.2009.79>
- Millikan, R. G. (1984). *Language, Thought, and Other Biological Categories*. MIT Press. <https://mitpress.mit.edu/books/language-thought-and-other-biological-categories>
- Millikan, R. G. (1993). *White Queen Psychology and Other Essays for Alice*. Bradford Books.
- Millikan, R. G. (2004). *Varieties of Meaning*. MIT Press.
- Neander, K. (2017). *A Mark of the Mental: In Defense of Informational Teleosemantics*. MIT Press.
- Owren, M. J., Rendall, D., & Ryan, M. J. (2010). Redefining animal signaling: Influence versus information in communication. *Biology & Philosophy*, 25(5), 755–780. <https://doi.org/10.1007/s10539-010-9224-4>



- Pandey, B., & Sarkar, S. (2017). How much a galaxy knows about its large-scale environment?: An information theoretic perspective. *Monthly Notices of the Royal Astronomical Society: Letters*, 467(1), L6–L10. <https://doi.org/10.1093/mnrasl/slw250>
- Piccinini, G. (2020). Nonnatural Mental Representation. In J. Smortchkova, K. Dołęga, & T. Schlicht (Eds.), *What are Mental Representations?* (pp. 254–286). Oxford University Press. <https://doi.org/10.1093/oso/9780190686673.003.0010>
- Piccinini, G., & Scarantino, A. (2011). Information processing, computation, and cognition. *Journal of Biological Physics*, 37(1), 1–38. <https://doi.org/10.1007/s10867-010-9195-3>
- Rathkopf, C. (2017a). Neural Information and the Problem of Objectivity. *Biology & Philosophy*, 32(3), 321–336. <https://doi.org/10.1007/s10539-017-9561-7>
- Rathkopf, C. (2017b). What kind of information is brain information? *Topoi*, 39, 95–102. <https://doi.org/10.1007/s11245-017-9512-6>
- Riley, J. R., Greggers, U., Smith, A. D., Reynolds, D. R., & Menzel, R. (2005). The flight paths of honeybees recruited by the waggle dance. *Nature*, 435(7039), 205–207. <https://doi.org/10.1038/nature03526>
- Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3-4), 379–423, 623–656. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press. <http://www.press.uillinois.edu/books/catalog/67qhn3ym9780252725463.html>
- Shea, N. (2018). *Representation in Cognitive Science*. Oxford University Press.
- Skyrms, B. (2010). *Signals: Evolution, Learning, and Information*. Oxford University Press.
- Woodward, J. (2003). *Making Things Happen*. Oxford University Press. <https://doi.org/10.1093/0195155270.001.0001>

Open Access

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

