

Consequences of a functional account of information

Stephen Francis Mann

the date of receipt and acceptance should be inserted later

Abstract This paper aims to establish several interconnected points. First, a particular interpretation of the mathematical definition of information, known as the causal interpretation, is supported largely by misunderstandings of the engineering context from which it was taken. A better interpretation, which makes the definition and quantification of information relative to the function of its user, is outlined. The first half of the paper is given over to introducing communication theory and its competing interpretations. The second half explores three consequences of the main thesis. First, a popular claim that the quantification of information in a signal is irrelevant for the meaning of that signal is exposed as fallacious. Second, a popular distinction between causal and semantic information is shown to be misleading, and I argue it should be replaced with a related distinction between natural and intentional signs. Finally, I argue that recent empirical work from microbiology and cognitive science drawing on resources of mathematical communication theory is best interpreted by the functional account. Overall, a functional approach is shown to be both theoretically and empirically well-supported.

Keywords Mathematical communication theory; Teleosemantics; Sender-receiver framework; Primitive content; Rate-distortion theory

1 Introduction

Several misconceptions about the application of information theory in natural science are widespread in philosophy. This paper deals with some of the core mistakes, demonstrating how they are mutually reinforcing and how they can be overturned. At the centre of the theoretical tangle is Dretske's interpretation of mathematical communication theory and concurrent definition of information. Here I promote a different account of information for natural science, one that makes it relative to a function or functions. This move allows us to solve problems and avoid misinterpretations of the mathematical concept of information.

The paper is structured as follows. Section 2 outlines two core assumptions of mathematical communication theory. To apply its mathematical tools elsewhere in natural science,

it will be necessary to dispense with these assumptions when investigation shifts to wider contexts. Section 3 then shows how Dretske failed to appropriately generalise communication theory for a wider setting. His account led to the contemporary causal interpretation of information, and I argue that an alternative functional interpretation is preferable. Following the first half of the paper, sections 4 and 5 explore, respectively, theoretical and empirical consequences of a functional account of information. Highlights include refuting the claim that quantification of information is irrelevant for meaning, rejecting a popular distinction between causal and semantic information, and showing that recent work in diverse areas of natural science accord well with a functional approach. Section 6 concludes.

2 Perspectives on communication theory

In order to locate the positive thesis, an overview of recent philosophical claims about the application of information theory across the natural sciences is in order.

Phenomena in several subdisciplines of biology and cognitive science recommend the use of information theoretic formalism. But differing assumptions entail different interpretations of formal results. Unfortunately, the special details of mathematical communication theory lead many to import one of its central tenets into areas where it does not belong. It is often claimed that Shannon and Weaver (1949) established that information theoretic formalism, in any domain, is irrelevant for the meaning of information transmitted. The claim is false but widely believed (Owren et al., 2010: 761) (Piccinini and Scarantino, 2011: 21) (Lombardi et al., 2015: 1989). It is false because Shannon and Weaver were not concerned with defining information in domains outside communication theory, so they could never have established such a claim. Their definition was relative to a particular framework. In order to apply the definition more widely, we must understand how to generalise the framework. Philosophical understanding of information in natural science is misshapen.

One particular detail stands out as a potential source of confusion. Mathematical communication theory (MCT) deals with symbols that stand for symbols. The encoded signal for which information is quantified represents a string of symbols whose meaning is irrelevant for this quantification. In other contexts, information can be quantified for signals that stand for things other than symbol strings. The engineering context is special, but its mathematics are general. A coherent account of information, and an understanding of the fruitful application of existing mathematical tools across natural science, results from rejecting the assumption that only signals standing for symbol strings can be quantified. This is the subject of the present essay.

The remainder of this section introduces MCT by detailing two components of the engineering framework that are often taken to be necessary to interpret the mathematics. Both were introduced by Shannon (1948) as part of his foundational text. First, the central transmitter-receiver model, which is applied in diverse scientific domains. Second, the engineering problem whose solution is the aim of communication theory has sometimes been taken as fundamental to the broader domain of information theory. The crux of the present essay is that contemporary philosophical interpretations of both of these factors are largely incorrect. In the remainder of this section I survey them in order.

2.1 The central model

In MCT, information is quantified as the extent to which the source message can be recovered at the target. To be transmitted, the source message is transformed into an encoded message. Information transmission is a function of statistical properties of the source message and the channel through which it is sent. The meaning of the source message – whatever its symbols represent or indicate – is irrelevant to measuring transmission. We might call this situation the “central model” of MCT.

As an example of communication, the central model is a rather peculiar case. Rarely does it apply exactly outside communications engineering. In order to apply information theory in other domains, the formalism has to be generalised. To this end, consider one of the special properties of messages in the central model. Because the symbols in the encoded message are transformations of symbols in the source message, encoded messages usually carry two meanings, one folded within the other. First, an instruction how to recover the original symbol string. Second, and as a result of the first, the meaning of the original symbol string, if it has one. It is often pointed out that the formalism of information theory is blind to the second meaning. While this is true, it neglects the possibility that formalism captures the first meaning, the instruction how to recover the original symbol string. Indeed I argue below that this is precisely what it quantifies. Moreover, there are communication systems whose messages are not encoded in this way. In these systems, the meanings of messages are quantified by information-theoretic formalism.

Since Bar-Hillel and Carnap (1953), philosophers have been told there is a deep divide between “Shannon information” – codes – and “semantic information” – what is expressed by codes. But whether or not the source message has a semantic meaning, the encoded message certainly does, and it is a meaning that is directly relevant to the quantification of information transmitted by it. Below in section 4 I demonstrate that the first meaning of an encoded message – the instruction (imperative) how to recover the primary message, or equivalently the information (indicative) about what it was – is a kind of primitive content familiar from signalling games. It is only because the central model is a very special kind of system that its relation to other forms of communication has been neglected by philosophers. In this way, contemporary positions on the use of information theory in natural science are inordinately pessimistic.

The central model has been applied in diverse ways in the natural sciences. There has been some difficulty establishing the conditions under which it is appropriate. Below in section 5 two examples are presented, and the justification for the application of the central model in those domains is examined.

2.2 The fundamental problem

Shannon described the problem of his art as “that of reproducing at one point either exactly or approximately a message selected at another point” (Shannon, 1948: 379). A similar lesson applies here. Just because the fundamental problem does not reappear exactly in natural science does not mean the formalism of information theory cannot be applied outside mathematical communication theory. Mathematical formalism, appropriately generalised, is never so rigorously context-bound.

To see how information theoretic formalism can be generalised beyond the fundamental problem of MCT, consider the qualifier “approximately” in Shannon’s quote above. Where exactitude is not required, the system may be optimised to transmit at a rate of ‘just enough’

information. But how can we determine how much is enough unless the measurement of information has relevance for what the receiver does with it? In other words, how do I know how many bits I need unless I know what actions those bits are helping me choose between, or which states of the world those bits are helping me infer? The cost of information loss is always measured relative to the goal that information transmission subserves. Cognitive science and microbiology are applying these ideas already (see below section 5). Philosophy of information needs to catch up.

The aim of this section was to pump intuitions against the received view of information theory and its application outside MCT. The next section carries those doubts which have hopefully been raised, and soothes them by providing an inclusive understanding of information.

3 Causal and functional interpretations

In this section I lay out the canonical interpretation of MCT in naturalistic philosophy. The interpretation is due primarily to Dretske, who developed a general application of the central model of MCT and corresponding general definition of information. Following this, in section 3.2 I outline an alternative account due primarily to Millikan. I introduce both accounts briefly, because my aim is to describe their motivations and characteristics before showing how the beneficial consequences of a functional account speak in its favour. The functional account I endorse was not originally proposed in relation to MCT. Nevertheless, we shall see in section 4 that the central model of communication theory has more in common with models inspired by the functional approach than has previously been supposed.

3.1 Dretske's causal interpretation

In this subsection I give an outline of the causal interpretation of information due mainly to Dretske (1981). I also describe one well-known problem for it: the reference class problem.

One of the key motivations behind Dretske's analysis of information was to define an objective, mind-independent resource that agents could use to make reliable inferences about the world beyond their sensory reach. Information had to be objective – definable independently of any given agent – because only then could it ground an explanation of the emergence of rational, conscious agents (Dretske, 1981: vii). Dretske's interpretation, which was to become the canonical account of causal information, made "information" extremely broad. To distil the narrower notion of semantic content he showed how more stringent conditions apply in the process of belief formation.

It is only a short step from MCT to Dretske's definition of information. He argued that both the mathematical tools and the central model that interprets them (described above in section 2.1) are universally applicable. Any medium through which correlations are borne is a channel. Any source of correlation – biological, artificial or inert – is a transmitter. And anything capable of interpreting that correlation – given a plausible construal of "interpretation" – is a receiver. For example, the nuclear interactions that take place in the sun generate electromagnetic radiation. This radiation tends to strike the sunward surface of the earth, and could therefore be used as information about the sun's position in the sky. Interference such as thick cloud cover distorts the signal, which entails that the amount of information received on the ground – and concurrently the accuracy with which an inference can be made – is reduced.

On this interpretation, information is quantified independently of any observer. The statistical probabilities governing the behaviour of the transmitter determine the information carried by the signal. These probabilities are objective chances. They are properties of events in the world, and are not defined relative to a potential or actual observer. This satisfies Dretske's criterion for a user-independent resource at the foundation of naturalist epistemology. The MCT channel is applied to situations where there is causal influence but no design. In the above example, the sun's production of electromagnetic radiation is not guided by earthlings' need to infer its position in the sky. This is what the label 'causal' connotes: information is defined with respect to objective statistical probabilities and chains of causal influence, not designed systems.

The desideratum of user-independence that Dretske felt was necessary for a principled definition of information opens the door to critique. A version of the reference class problem, familiar from philosophy of probability (Hájek, 2007), was first raised against Dretske in Harman's early commentary (see p.72 of Dretske (1983) for the commentary and p.84 for Dretske's reply). The objection runs as follows. A given instance of an event can only be ascribed a probability with respect to a wider class of events of the same kind. Tokens, in other words, have probabilities only in virtue of the types to which they belong. But token events do not belong unequivocally to any given class of events. In most cases, there is no principled way to choose the wider set to which a given event belongs. As a result no definite statistics govern a particular instance of the source event in Dretske's model, so no definite quantity of information can be obtained for the consequent signal.

A comprehensive defence of causal information in light of the reference class problem might take up a whole paper (see Kraemer (2015) for example). Here we need only take account of it as part of the motivation for an alternative, functional approach. User-independence prompts the reference class problem, and as we shall now see, construing information as relative to a user helps solve it.

3.2 A functional interpretation

In this section a functional interpretation of information is outlined.¹ I describe how user-relativity provides a solution to the reference class problem. Theoretical consequences are discussed in section 4, while examples from empirical science are considered in section 5.

The label 'functional' encompasses two aspects worthy of immediate note. The first is the user-relative nature of the definition, which is the subject of this subsection. The second is the pragmatic aspect of information, which is often emphasised in the sciences, as opposed to its epistemic or 'inferential' nature which philosophy puts to the fore. Section 4 advocates a return to theory that focuses on information as a practical resource. In sum, a focus on the user solves the reference class problem, and a focus on the user's function provides a general foundation for information in natural science.

The functional interpretation was prefigured by Dretske himself. Part of Dretske's definition concedes that the information carried by a signal is sometimes determined with respect to its receiver. Some signals only carry information if their receivers have sufficient

¹ For a history of approaches to information based on purposive behaviour, see Adams (2003: §3). More recent work includes Bergstrom and Rosvall (2011), which offers a user-relative definition of genetic information, and Rathkopf's account of neural information, which defends a function-relative definition against the worry that such an approach might threaten scientific objectivity (Rathkopf, 2017). See also Dennett (2017: §6) and Fresco et al. (issue). Space precludes a comparison between these works and the account discussed here, but I suspect they are all broadly compatible.

'background knowledge' to interpret them. But Dretske felt background knowledge had ultimately to be cashed out in terms of information. In contrast, the functional approach takes a general kind of 'background knowledge' to be antecedent to information use. Supposing we have an account of what would constitute a user of information, we can define information relative to it. As a result, we can extend what Dretske considers a special case to all cases. We can consistently demand that information be definitionally relative to a user's background 'knowledge'.

This is the line taken by Scarantino (2015), who describes a solution to the reference class problem (and several other problems) as a result of a user-relative definition of information. Very roughly, what an agent already 'believes' will determine the quantity of information it can receive from a given signal. Here, 'believes' is in scare quotes because it is to be read functionally or dispositionally: its 'beliefs' are the propositions that must be true for its behaviour is to be successful. Scarantino's less provocative label for the same thing is 'background data'. When behaviour changes upon detection of a stimulus, we can model that stimulus as an information source which, combined with the agent's background data, produces that behavioural change. For example, if a rat is able to learn that food will be served at the ring of a bell, its background data will contain a proposition like 'the bell ringing entails food will almost certainly be served soon'. The information carried by the ring of the bell is quantified by how much more probable the occurrence of food now is – which is determined by the 'almost certainly' part of the phrase above.² According to this framework, in order for a stimulus to carry information, the agent must already be set up to respond to it – in other words, its background data must already contain a hypothesis regarding the statistics governing the stimulus.

To reiterate: any agent, the success of whose behaviour depends on some external condition, could be described as a potential user of information about that condition.

The statistics defined by an agent's background data offer a principled solution to the reference class problem. If we can show that background data is determined by an agent's behaviour, then we can quantify the information carried by a stimulus with respect to that agent. Like the weight of an object considered with respect to different planets, the information carried by a signal – its 'weight of evidence' – varies depending on the consumer of that signal. Agents need not have full belief-desire psychology for this descriptive strategy to work. They need not represent background data to themselves. Scarantino's framework, which falls into the broad class of cognitive models labelled 'Bayesian', applies to simple biological systems with the ability to respond to environmental stimuli. This is in part because it is intended as an extension of Millikan's approach (Millikan, 2013a), which is itself a response to Dretske's account (Dretske, 1988: §3.2). All of these works are aimed at describing information consumed by simple biological systems, not just sophisticated cognitive agents.

This functional account accords well with statistical decision theory, which provides strategies for agents depending on their goals and the information available to them. As Stegmann (2013: p.7, Box 1) points out, information in this framework is quantified relative to the agent in question. Ecological models in the Bayesian paradigm define and measure information with respect to an agent's prior and posterior distribution over relevant states (see for example Dall et al. (2005: p. 189 Box 1)); this is what Scarantino seeks to capture. Although classical decision theory typically construes agents as rational, the same mathematics applies to strategies designed by evolution. In biological settings, we can explain organisms' responses to stimuli based on the information carried by those stimuli. But this explanation

² For more on informational readings of classical conditioning, see Rescorla (1988) and Gallistel (2003).

is available only because of the learned or evolved response of the organism. Quantifying information requires determining statistical correlations that held during the learning period. This is the spirit of Millikan's original solution to the reference class problem (Millikan, 2013a: §5.3).

On this approach, the reference class problem is avoided. A principled reference class exists: the tokens of the stimulus encountered during the learning period that contributed to the strategy the agent currently employs. This, I take it, is how to cash out 'background data' for agents without full-fledged representational systems. There remain problems for such an account that I will not address – for example, which interactions in the learning period count as stimuli, or how the learning period is itself determined. But I hope the spirit of the account is clear. Functional behaviour is defined first, and information is defined in terms of its contribution to the success of that behaviour.³

The next two sections discuss consequences of such an account. Section 4 deals with theoretical consequences, demonstrating how MCT employs a definition of information that is derivable from this one, and how the popular 'irrelevance claim' is fallacious. Section 5 demonstrates recent empirical work that makes use of MCT and how it can be fruitfully interpreted in light of this functional account.

4 Theoretical consequences: the irrelevance claim and the causal/semantic distinction

In this section I argue for two claims. First, mathematical communication theory is not irrelevant for semantic concepts of information. Although Dretske was wrong in the specifics of his approach, communication theory can be leveraged to understand content in signalling systems. This undermines a popular distinction between causal and semantic information. Second and consequently, the explanatory work currently assigned to the causal/semantic distinction is best performed by a more clearly understood distinction between natural and intentional signs. Intentional signs are transmitted and received by codesigned entities, while natural signs are received by an entity not codesigned with the information source. The conceptual differences between these two categories, as well as the range of formal tools required to analyse them, are well understood and well supported.

4.1 Overturning the irrelevance claim

The irrelevance claim is the claim that formal measures from MCT have no relevance for the meaning or content of signals. I argue against it by applying the recent concept of subpersonal content. I demonstrate that signals in the central MCT model possess subpersonal content by definition, and it is this that is quantified by informational equations.⁴ To approach the argument, I beg the reader's patience while a little background is put in place, introducing the notion of subpersonal content and its significance for the issues at hand.⁵

A growing trend in naturalistic epistemology is the use of a concept of content that does not require personal-level intentional states. Its value lies in its explanatory role describing

³ The 'success semantics' account of mental content has a similar structure (Whyte, 1990).

⁴ See also Mann (2018: 10-12).

⁵ The irrelevance claim has been challenged by others, but it seems tough to overturn: "By the time of our Third London Symposium on Information Theory in 1955, it had become something of an accepted saying that 'information theory has nothing to do with meaning'. The time seemed ripe to question this hardening dogma..." (MacKay, 1969: 79)

the behaviour of organisms and artificial devices too simple to be ascribed personalities (Millikan, 1995; Shea et al., 2017). The classic example of this move is Skyrms's account of signals (Skyrms, 2010), which borrowed formal tools from Lewis (1969). Lewis applied a game theory framework to study the behaviour of rational actors, extracting a notion of content from the dynamics of behaviour observed in such games. Skyrms demonstrated that rationality on behalf of the actors is unnecessary. The same notion of content – the same explanations and descriptions of behaviour of the players – can be put to work when agents in the game act in accordance with evolutionary design rather than rational decision. As with the theoretical approach spearheaded by Millikan and Shea, this move can be regarded as a demonstration that concepts previously developed for the intentional level are applicable at the design level too.

Consider the following similarity between the formal concept of information and sub-personal content. One of the important aspects of content, discussed by Millikan, Lewis, Skyrms and others, is its dual aspect of indicative and imperative conditions. Signals – intentional signs – can say both how things are and what is to be done.⁶ When it comes to information we tend to consider only the indicative aspect of signals. In turn, we favour an epistemic interpretation of its use, on which information is an inference-supporting resource, independent of how an agent might act in response to it. The agent may be actual or hypothetical, but it is their actual or possible knowing that lends an indicative flavour to “information”. I contend that once we focus not just on knowing but on actual or possible *doing* we regain the instructional aspect of signals, thus moving closer to the contemporary understanding of content.⁷

In terms of subpersonal content, an encoded signal in the central model of communication theory is *primitive*. Primitive content denotes a signal that is equally indicative and imperative. We can see that this is true of the encoded signal in the central model by overlaying a sender-receiver model on the inner portion of the Shannon system, taking the sender to be the encoding transmitter and the receiver to be the decoding receiver. The signal is then the stream of encoded symbols. Sender and receiver are dealing with encoded information whose ‘meaning’ is the primary message. The information can equally be seen as telling the decoder what the source message is (informative) and telling the decoder which target message to construct (instructional).

A test for primitivity is deliberation (Lewis, 1969: 144) (Huttegger, 2007: 410). Where receivers are permitted to deliberate over the action they will perform, the signal they receive has a more indicative flavour. In contrast, where senders are permitted to deliberate over what signal to send, it seems they are telling the receiver what to do. (If both are permitted to deliberate, the correlational link between world and act is in danger of being destroyed, and communication breaks down.) In the present case, neither encoder nor decoder deliberates. Neither draws on information outside the channel through which the signal in question is flowing. The encoder takes the message as input and produces a stream of code. The decoder takes that code as input and produces a message. Since the same primary message cannot give rise to a different code, the encoder does not deliberate. And since the same code cannot be translated into different messages, the decoder does not deliberate. Finally, since neither deliberates, the code is an instance of primitive content.

⁶ In the terminology of traditional philosophy of language, signs can have both kinds of ‘direction of fit’ at once (Millikan, 1995).

⁷ It is not immediately clear how to interpret the formalism, but here is a suggestion: while 1 bit of information allows the receiver to infer one out of two equiprobable states of the world, perhaps 1 bit of instruction allows the receiver to choose one out of two equiprobable acts.

The previous two paragraphs showed that signals in the central model of MCT can be consistently interpreted as having semantic content. It is worth noting this holds regardless of the meaning of the source message. Indeed, the source “message” need not be a string of symbols. It need only be an element or sequence selected from a set. Similarly, the target need only be an element or sequence drawn from a set, and it need not be the same set as the source.⁸ As pointed out in section 2.2, informational equations can be applied outside the restricted context of the fundamental problem.

Within the central model, encoded messages possess primitive content. Compare Piccinini and Scarantino (2011: 19): “Shannon’s messages need not have semantic content at all – they need not stand for anything.” (Compare also Owren et al. (2010: 761).) Taken literally, this is false. An encoded message in the central model must stand for its source message, otherwise there can be no definition of information rate. The authors might reply that what they meant to say is that the meaning of the *source* message is irrelevant to the quantification of information transmission. It could be a string of meaningless symbols and transmission rate would not change. But this latter point is repeatedly conflated with the much stronger and entirely unsupported claim that “Shannon information” is irrelevant for meaning *in all domains in which information can be quantified*. It is by failing to appreciate the special nature of the central model that the claim of universal irrelevance gains traction.

Primitive content is prevalent in simple systems, which is why it was christened “primitive” by Harms (2004). Where what matters is coordinated behaviour, being told *that* another agent is performing act A is equivalent to being told *to* perform act B. Studies of simple agents in signalling games demonstrate that the indicative and imperative aspects of subpersonal content – the informative and instructional aspects of the MCT concept of information – are useful concepts to apply at the level of the design stance. The causal interpretation encourages inordinate emphasis on the epistemic. Naturalistic intentionality is better off grounded in function. Both informative (indicative, descriptive) and instructional (imperative, prescriptive) aspects are of equal significance.⁹

To sum up: within the informational paradigm, there is something that looks like instruction or can fruitfully be interpreted as such. Information and instruction look like indicative and imperative content. We already have a comprehensive account of content – the teleosemantic/evolutionary game theory approach whose unification has recently been argued by Artiga (2016: 494-6) – that endorses amalgamating indicative and imperative content for simple systems. The bottom-up approach of signalling systems coincides with the top-down approach of teleosemantics. We should embrace their coherence. In contrast, the causal interpretation of information has no notion of user action. It embodies an epistemic approach that focuses on knowing rather than doing. The next subsection uses these considerations to argue that the popular distinction between causal and semantic information is misleading, and should be replaced by a distinction between natural and intentional signs.

⁸ Shannon (1959: 326-7) points out that source and target sets need not be identical, nor even ‘symbols’ in the familiar sense. Lean (2016: 239-40) makes a similar point, forging a path for the adoption of informational models in other domains.

⁹ A recent trend seeks to distinguish two concepts I treat as equivalent. The distinction, advocated by several authors including Price (2008: §5), Hutto and Myin (2013: 67), Rescorla (2013) (who cites Burge (2010) as inspiration) and Lean (2014), runs as follows. Simple signalling systems carry information in the guise of *reliable correlation* (“functional isomorphism”, “Shannon information”) – tokens that correspond to worldly states in a manner sufficient for successful behaviour. But correlational information is to be distinguished from the much richer notion of *content*, which is characterised by truth conditions. There is far more to say about this distinction and its motivations than can be addressed here. For a defence of the use of semantic concepts in simple signalling systems, see Millikan (2013b). For some remarks on the term “Shannon information” see below, page 11.

4.2 Natural signs and intentional signs

In this subsection I demonstrate that the natural/intentional distinction is well placed to do the explanatory work typically ascribed to the causal/semantic distinction. ‘Natural information’ denotes information whose sender is not codesigned with its receiver. A receiver benefits from learning how to respond to the stimulus, but the source of the stimulus does not benefit – either because it has incompatible interests with the receiver or because is not an agent at all. In contrast, ‘intentional information’ describes a cooperative relationship between sender and receiver, where both benefit from coordinating their behaviour with the use of a signal. I shall argue that the natural/intentional distinction is useful and accords with both theory and practice in natural science. Meanwhile, the causal/semantic distinction is inspired by a confusion about MCT, and obscures fruitful relationships between models of communication and strategic behaviour.

The causal/semantic distinction (Piccinini and Scarantino, 2011: §§4.1-2) (Godfrey-Smith and Sterelny, 2016: §§2-3) has at least two sources. It is firstly a mutated form of an earlier distinction between natural and intentional meaning, which may be traced back at least to Brentano and found its clearest statement in Grice (1957). Prompted by Dretske (1981, 1988) the distinction took centre stage in the teleosemantic literature of the 90’s (Millikan, 2001). The original distinction is still hard at work in Millikan’s teleosemantics (Millikan, 2017: §§11-12), but its mutated form is misleading. A second source is Bar-Hillel and Carnap’s clarification of “information” as it appears in MCT. They distinguished the mathematical quantity from the semantic notion which is of interest to philosophers (Bar-Hillel and Carnap, 1953). Dretske compared Grice’s approach, as well as that of Bar-Hillel and Carnap, to his own project (Dretske, 1981: pp. 241-2, n.1 and n.10). Soon after, the “still imperfectly understood” distinction was cited by Dennett (1983: p. 344 col. 2) and picked up by Krebs and Dawkins (1984: §§4.1-2), whence it found its way into the behavioural ecology literature and prompted ongoing scepticism about the use of information theory in the study of animal signalling (Pfeifer, 2006; Owren et al., 2010; Sarkar, 2013).

The causal/semantic distinction is predicated on taking a causal interpretation of MCT along with the irrelevance claim. Together, they entail that whatever semantic information is, it must be something richer than causal information, something that requires a different formal framework.¹⁰ This distinction has become so widely accepted that the Stanford Encyclopedia entry “Biological Information” is currently organised around it (Godfrey-Smith and Sterelny, 2016). I have already argued against both premises of the distinction: section 3 offered a better interpretation of MCT while section 4.1 overturned the irrelevance claim. I will now demonstrate that the job vacated by the redundant causal/semantic distinction is best performed by the natural/intentional distinction.

Natural and intentional signs lie on a common spectrum. Just as there can be degrees of adaptation leading from disposition to behaviour, there are degrees of coadaptation from joint disposition to joint behaviour. A sign that mediates an interaction becomes more ‘intentional’, on this definition, when production of and response to the sign become coadapted behaviours. The sign is more ‘natural’ as production and response are less coadapted. Behavioural ecology affirms this idea with a useful distinction between signals and cues. Signals are typically defined in terms of coadapted behaviour. Cues can be environmental stimuli or behaviour performed by other organisms, but they are opportunistically exploited

¹⁰ Due to space constraints I neglect several possible positions. For example, it is possible to accept the causal interpretation of information and deny that anything further is needed for semantic information; see Skyrms (2010: §3). I also ignore Grice’s term “non-natural meaning” to prevent confusion.

by receivers. Sometimes cues become signals, when the receiver's behaviour turns out to be beneficial to the sender. This process is called ritualisation, because the properties of the cue tend to become 'ritualised' – made more salient, easier to recognise, or repeated several times – in order to promote the receiver's behavioural response.¹¹

A natural sign has no function *qua* natural sign. Whatever 'sends' a natural sign, if it has a function, is not the same function as that for which the natural sign is used. So for example, a signal sent between two cooperating entities is an intentional sign for them, but is a natural sign for an eavesdropper.¹² The contemporary view seems to have confused natural signs for "causal information", which prompts the belief that informational quantification can only measure natural signs. Given that anything could in principle become a natural sign if there was an agent who could make use of it, causal information is thought to be an incredibly broad resource, and as a result virtually useless for biological theorising.

Contemporary scholars, I suggest, often unknowingly aim for the natural/intentional distinction, especially when they invoke "Shannon information". The popular claim is that Shannon information captures statistical correlations, and this alone cannot distinguish between natural and intentional signs – between cues and signals (Godfrey-Smith and Sterelny, 2016: §2) (Owren et al., 2010: 772-3). However, the term "Shannon information" seems to promote confusion. It is true that the information measure employed by statistical decision theory is typically applied to natural signs and not intentional signs. It is also true that this measure is derived from Shannon's work (Shannon, 1948) which is itself a continuation of Hartley (1928). But communication theory has other formal tools beyond measures of correlation. It has other commitments besides the condition that signals covary with their sources. Communication theory is not simply statistical decision theory, though it is closely related. The mistake is likely promulgated by the association of Shannon's name with the quantity used in decision theory. It might be best to retire the misleading label.

Overall in this section, we have seen grounds to reject some popular positions within the philosophy of information. At the very least, the central model of MCT is apt for analysis in terms of one of the meanings of its signals. The next section goes further by examining two recent applications of MCT in natural science. Although the earliest theoretical work was concerned with countering noise in order to eradicate error completely, subsequent theory generalised that problem to one where a certain level of error is tolerable. This already makes its models more amenable to biological application.

5 Empirical consequences: rate-distortion theory in natural science

In this section, I show how two case studies bring out a wider lesson recommending function as a foundation for the application of information-theoretic concepts in cognitive science and biology. I discuss the case studies in the following two subsections, before describing how the functional account helps to account for these applications of information theory.

¹¹ I omit a third category, manipulations, which are influential behaviours performed by senders to the detriment of receivers. These may display the same 'ritualised' qualities described for signals. The theoretical status of manipulations is still in dispute, with some arguing they should be included in the definition of signals (Owren et al., 2010). I resist that categorisation because I see coadaptation (or more broadly, codesign) as central to the mathematical and conceptual tools we use to analyse the varieties of information. On the other hand, Owren et al. (2010) see information as a deeply problematic concept that should be left out of animal communication studies altogether. See Mann (2018) for a fuller discussion.

¹² It is this difference that Millikan (2013a) leverages to analyse the correctness conditions of intentional signs in terms of their function. Natural signs, by definition, have no correctness conditions; they are neither true nor false.

To anticipate, a functional account offers a clear interpretation of the cost of error, which is central to the case studies.

5.1 Microbiology

Many single-celled organisms can sense chemical changes in their surroundings. By navigating along gradients of changing density they can find food or avoid toxins. One species, *Dictyostelium discoideum*, uses this process of chemotaxis to coordinate mass response to a lack of nutrients. When food is scarce it is beneficial to pool resources by aggregating. *D. discoideum* cells seek each other out by alternately releasing waves of chemicals and moving in the direction of greatest concentration. When enough cells aggregate, a fruiting body forms, which helps propagate spores into a more favourable environment. These become the next generation of cells, once again adopting an individual lifestyle.

During the aggregation phase, individual cells face an informational problem. They need to determine the best direction in which to travel, and they need to be sensitive to external changes in order to do this. The metabolic cost of sensitivity to fine changes in gradient impedes perfect behaviour. Like many other living things, *D. discoideum* must strike a balance. It must optimise its behaviour relative to informational constraints and the requirements of behavioural accuracy. Fortunately, rate-distortion theory is designed to analyse such trade-offs. At the heart of the theory is a cost function describing the penalty for misinterpreting a signal. Depending on the cost an agent is willing to incur, it is permitted to obtain smaller amounts of information. Because each situation – across engineering, computer science, cognitive science and biology – is likely to involve different trade-offs between information and cost, the rate-distortion function describing optimal behaviour is usually calculated anew each time.

In recognition of the problem facing *D. discoideum*, Iglesias (2016) describes how rate-distortion theory could be applied to explain and predict optimal behaviour. He uses a non-standard application of the central model to describe the cell's predicament. The signal is not the external chemical gradient. Instead, it is the hypothetical decision located between the cell's receptors (whose state corresponds more or less exactly to the immediate gradient) and the cell's behaviour. The central model is applied entirely within a single cell. Mathematically the approach is acceptable, since a channel need be nothing more than a probabilistic connection between two pieces of behaviour. Given this way of modelling the situation, an appropriate optimisation function describes the amount of information required by a migrating cell to successfully reach its target, which is equivalent to how well-correlated the cell's decision should be with its external receptors. How can we estimate information rate and cost schedule in order to derive such a function? Iglesias takes a function that measures angular deviation from the direction the cell is supposed to be moving towards. Cost increases as the angle between the cell's movement and the true direction of aggregation increases.

Iglesias provides a plausible starting point for the application of information theory to optimisation problems of this kind. However, some of the details are as yet unjustified. It is not clear how sharply cost should grow as angular deviation increases. Nor is it clear how this function might change as the cell approaches its target. Presumably the cell needs to be more sensitive to information the further it is from the goal. As it gets closer, its internal bias could override transient changes of chemical gradient that would otherwise send it in the wrong direction. Iglesias considers the possibility of internal bias, but only to demonstrate how it affects the mathematics in an idealised case. Empirical work is required to determine how much and what type of bias develops during chemotaxis. This entails a methodologi-

cal problem: both bias and cost function must be derived from behaviour, but both of them are unknown or only broadly guessable at the outset. Perhaps parametric models describing both functions at once can be employed to generate testable hypotheses. These are typical methodological issues faced when fitting a model to reality. The application of information theory in microbiology is in its early days, but signs are positive it can provide a real contribution.

The next subsection describes the application of rate-distortion theory in a rather different domain, namely human perception.

5.2 Perception

We just saw an application of rate-distortion theory for which the output of information transmission was a piece of behaviour, namely directional orientation and movement towards a goal. Sims (2016) applies rate-distortion theory in a domain closer to its original home, in which the output of information transmission is another piece of information.

Sims discusses two experiments. The first examines how accurately human subjects can categorise straight lines by length (“absolute identification”). The second determines how accurately subjects can choose the longest of two lines seen one after the other (“perceptual working memory”). Accuracy in each of these tasks is impeded by the information capacity of perceptual processing. Sims applies rate-distortion theory to estimate the optimum rate given the cost of inaccuracy. He places emphasis on the lack of a known cost function, and the methods by which we can estimate one. As a result of applying the theory to both experimental procedures, Sims purportedly derives new insights into human perceptual performance (Sims, 2016: p. 185 col. 2 and p. 190 col. 2).

Consider the absolute identification task (taken from Rouder et al. (2004)). A set of lines of N different lengths were presented randomly. Subjects were asked to choose the category, from 1 to N , in which the line belonged. After each choice the subject was informed whether or not they were correct. The mapping of the central model onto perception is more intuitive than the microbiological case. The source is the perceptual stimuli – lines of differing lengths – and the output is the subject’s response. Encoding and decoding take place within the subject’s perceptual system, and the channel capacity is inferred from the probabilities of correct responses as the number of possible lines increases. Subjects seem to reach a point at which their performance cannot improve indefinitely (Sims, 2016: 185-6), implying capacity is limited and rate-distortion theory can be fruitfully applied.

What are the cost functions constraining human perceptual performance? Sims admits they are largely unknown. But he offers resources for estimating them from performance data. His figure 5 (Sims, 2016: 186) depicts three different models of increasing fit with the data, corresponding to three different cost functions of increasing complexity. The final and most accurate cost function accords with existing ideas about perceptual “anchors” used by subjects to generate best guesses. By presenting a sequence of potential models, Sims advocates something like the following methodology. We can use rate-distortion theory to derive increasingly accurate estimations of the cost function guiding perceptual tasks, while at the same time providing hypotheses as to *why* those cost functions should be at work rather than some other. One of the problems with this approach is that the cost function is not the only unknown. In the second study, performance varies with the subject’s implicit estimation of source statistics (Sims, 2016: p. 191 col. 2). The experimenter must use performance data to infer both the subjective cost function and the subjective statistics. Similarly with the microbiology case, we have two unknown parameters and only one set of data for inferring their

values. As above, the situation is not insurmountable. Sims details methods for inferring appropriate models by using empirical data together with reasonable hypotheses.

Overall, Sims sees rate-distortion theory as a tool to investigate the processing capabilities of biological systems (Sims, 2016: p. 193 col. 1). Though he deals with what are essentially informational outputs – the responses of test subjects – he emphasises the generality of goals subserved by information processes (Sims, 2016: p. 193 col. 1): “The objective for biological information processing is not (merely) the communication of information, but rather the minimization of relevant costs. Information is simply a means to an end.” This approach accords with the interpretation of Iglesias presented above.

5.3 A functional account of information provides a clear interpretation of cost

We have seen two applications of rate-distortion theory in two rather different domains. Iglesias and Sims analyse biological signals with unavoidable reference to their meaning. This practice is more easily explicable on the functional view than the causal approach. One way to interpret cost is in terms of biological function. While the causal interpretation has been criticised as too broad and failing to prioritise useful information over idle correlations,¹³ the functional interpretation explicitly considers the cost of inaccuracy that is a consequence of reduced information rate. Here, information rate gets its significance from the magnitude of benefit it induces. Functional information is useful *ex hypothesi*, but there are principled bounds on its utility, as with any other resource.

In communications engineering, rate-distortion curves can be interpreted one of two ways. Suppose you know the maximum cost of inaccuracy you are willing to incur. Then the curve tells you the minimum information rate you need to transmit at in order not to exceed that cost. Alternatively, suppose you know the maximum rate you are able to transmit at. Then the curve tells you the minimum cost you can hope to incur. In biology, however, it seems cost will always come first. Obtaining information is a strategy for reaching a goal. The metabolic resources invested in gathering and transmitting information depend on how much you need, which is determined by a cost schedule covering the many ways of failing to achieve the goal. The cost of failure is then traded off against metabolic cost.

Increasing information rate in a communication system plausibly imposes metabolic costs. The situation is a special case of behavioural optimisation. Here a rate-distortion curve is the correct model to describe the relationship between improvement and metabolic cost, because the means of improvement is information transmission. Mathematical tools used to describe this kind of optimisation are taken directly from MCT. One significant change is that instead of state space being a message in a lexicon it is a biologically relevant state of affairs. The signal contains information about that state of affairs, and quantifying that information is a crucial aspect of explaining optimal behaviour. The application of information theory in biology and cognitive science concerns optimisation with respect to the statistical properties of both source and target, which need not be (and in the biological case hardly ever are) messages in well-defined lexicons. The same sentiment is echoed in recent work applying mathematical tools to molecular communication:

Signals are being regularly transmitted within and between individual cells and microorganisms. These signals may not be sending packets of data in the conventional

¹³ In particular by Rathkopf (2017), who advocates a relativist approach to neural information. Rathkopf also criticises the overly permissive notion of Shannon information as being out of step with the mathematical definition of information at work in engineering.

communication sense, but nevertheless they enable conventional communication applications such as sensing, coordination, and control. Thus, we can adapt conventional communication engineering theory and techniques to study these signaling mechanisms and understand how they work. (Noel et al., 2017: 1)

One final point about biological cost is in order. Ultimately, the costs that shape the behaviours described above are provided by natural selection. Environmental pressures determine optimal behaviour. Functions like microbe motility and perceptual categorisation are not performed for their own sake. They contribute to the survival and reproduction of whichever entity or entities are under selection. We should therefore expect the cost function of an individual piece of behaviour to derive from cost functions that govern selective pressures. Donaldson-Matasci et al. (2010) describe evolutionary fitness in informational terms. In an analysis deriving ultimately from Kelly's interpretation of information rate (Kelly, 1956), the fitness penalty of failing to heed information is bounded by the quantity of that information. Frank (2012) offers an informational interpretation of evolutionary fitness that seems to accord with this view. For simple models at least, our concepts of information and fitness cost are deeply entangled. When fitness is interpreted as growth rate, it can be measured with a unit that is commensurable with information units (Donaldson-Matasci et al., 2010: p. 228 col. 2), of which the 'bit' is a familiar example. And when cost functions are interpreted as fitness penalties, their proper unit of measure is the same as that of information. Rate-distortion curves can then be interpreted directly as the contribution to fitness afforded by information transmission. To apply these ideas more concretely, such as to the work of Iglesias, would require an understanding of how individual behaviours contribute quantitatively to the fitness of organisms.

In sum, a functional interpretation of information accords well with a growing trend that seeks to understand the optimisation of natural communication systems. In contrast to the causal interpretation, biological information cannot be pulled apart from the costs incurred to handle it and the benefits attained by using it. Information allows a principled redistribution of physical resources, entailing optimised behaviour that contributes to downstream functions and, eventually, evolutionary fitness.

6 Recapitulation

Four connected points have been raised in the paper.

1. The concept of "information" as defined in communication theory can be interpreted as relative to user function (section 3). Although the causal interpretation due primarily to Dretske is currently orthodox, it is supported by a misinterpretation of communication theory.
2. The claim that information is irrelevant for meaning in every domain in which it can be quantified is mistaken (section 4.1).
3. Points 1 and 2 together encourage us to replace the causal/semantic distinction with the natural/intentional distinction (section 4.2).
4. Contemporary uses of MCT in two diverse areas of natural science are well-interpreted on a functional account (section 5).

I hope to have shed light on one strand of a web of theoretical and empirical work organised around the vexed concept of information. The bottleneck between communication theory as

a mathematical and engineering discipline, and philosophical interpretations of “information” in natural science, is distressingly narrow. By exposing some pernicious misconceptions we can pave the way for a principled understanding of a naturalistically respectable concept of information that can do useful work in many scientific domains.

Acknowledgements Thanks to Ron Planer, two anonymous referees, and the editors for comments. Thanks also to Manolo Martínez for pointing me in the direction of rate-distortion theory. This research is supported by an Australian Government Research Training Program (RTP) Scholarship and Australian Research Council Laureate Fellowship Grant FL130100141.

References

- Adams, F. (2003). The Informational Turn in Philosophy. *Minds and Machines*, 13(4):471–501.
- Artiga, M. (2016). Teleosemantic modeling of cognitive representations. *Biology & Philosophy*, 31(4):483–505.
- Bar-Hillel, Y. and Carnap, R. (1953). Semantic Information. *The British Journal for the Philosophy of Science*, 4(14):147–157.
- Bergstrom, C. T. and Rosvall, M. (2011). The transmission sense of information. *Biology & Philosophy*, 26(2):159–176.
- Burge, T. (2010). *Origins of Objectivity*. Oxford University Press, Oxford.
- Dall, S. R. X., Giraldeau, L.-A., Olsson, O., McNamara, J. M., and Stephens, D. W. (2005). Information and its use by animals in evolutionary ecology. *Trends in Ecology & Evolution*, 20(4):187–193.
- Dennett, D. C. (1983). Intentional systems in cognitive ethology: The “Panglossian paradigm” defended. *Behavioral and Brain Sciences*, 6(3):343–355.
- Dennett, D. C. (2017). *From Bacteria to Bach and Back: The Evolution of Minds*. Penguin UK.
- Donaldson-Matasci, M. C., Bergstrom, C. T., and Lachmann, M. (2010). The fitness value of information. *Oikos*, 119(2):219–230.
- Dretske, F. (1981). *Knowledge and the Flow of Information*. MIT Press, Cambridge, Mass.
- Dretske, F. (1983). Précis of Knowledge and the Flow of Information. *Behavioral and Brain Sciences*, 6(1):55–90.
- Dretske, F. (1988). *Explaining Behavior: Reasons in a World of Causes*. MIT Press, Cambridge, Mass.
- Frank, S. A. (2012). Natural selection. V. How to read the fundamental equations of evolutionary change in terms of information theory. *Journal of Evolutionary Biology*, 25(12):2377–2396.
- Fresco, N., Jablonka, E., and Ginsburg, S. (this issue). Functional Information: A graded taxonomy of difference makers. *Review of Philosophy and Psychology*.
- Gallistel, C. R. (2003). Conditioning from an information processing perspective. *Behavioural Processes*, 62(1–3):89–101.
- Godfrey-Smith, P. and Sterelny, K. (2016). Biological Information. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Summer 2016 edition.
- Grice, H. P. (1957). Meaning. *The Philosophical Review*, 66(3):377–388.
- Hájek, A. (2007). The reference class problem is your problem too. *Synthese*, 156(3):563–585.

- Harms, W. F. (2004). Primitive content, translation, and the emergence of meaning in animal communication. In Oller, D. K. and Griebel, U., editors, *Evolution of Communication Systems: A Comparative Approach*, pages 31–48. MIT Press, Cambridge, Mass.
- Hartley, R. V. L. (1928). Transmission of Information. *Bell System Technical Journal*, 7(3):535–563.
- Huttegger, S. M. (2007). Evolutionary Explanations of Indicatives and Imperatives. *Erkenntnis*, 66(3):409–436.
- Hutto, D. D. and Myin, E. (2013). *Radicalizing Enactivism: Basic Minds Without Content*. MIT Press, Cambridge, Mass.
- Iglesias, P. A. (2016). The Use of Rate Distortion Theory to Evaluate Biological Signaling Pathways. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 2(1):31–39.
- Kelly, J. L. (1956). A New Interpretation of Information Rate. *Bell System Technical Journal*, 35(4):917–926.
- Kraemer, D. M. (2015). Natural probabilistic information. *Synthese*, 192(9):2901–2919.
- Krebs, J. R. and Dawkins, R. (1984). Animal signals: Mind-reading and manipulation. In Krebs, J. R. and Davies, N. B., editors, *Behavioural Ecology: An Evolutionary Approach*, pages 380–402. Blackwell Scientific, Oxford, 2nd edition.
- Lean, O. M. (2014). Getting the most out of Shannon information. *Biology & Philosophy*, 29(3):395–413.
- Lean, O. M. (2016). *Biological Information*. PhD thesis, University of Bristol, Bristol, UK.
- Lewis, D. (1969). *Convention: A Philosophical Study*. Blackwell, Oxford.
- Lombardi, O., Holik, F., and Vanni, L. (2015). What is Shannon information? *Synthese*, 193(7):1983–2012.
- MacKay, D. M. (1969). *Information, Mechanism and Meaning*. M.I.T. Press, Cambridge, Mass.
- Mann, S. F. (2018). Attribution of Information in Animal Interaction. *Biological Theory*, pages 1–16.
- Millikan, R. G. (1995). Pushmi-Pullyu Representations. *Philosophical Perspectives*, 9:185–200.
- Millikan, R. G. (2001). What has Natural Information to do with Intentional Representation? *Royal Institute of Philosophy Supplements*, 49:105–125.
- Millikan, R. G. (2013a). Natural information, intentional signs and animal communication. In Stegmann, U. E., editor, *Animal Communication Theory*, pages 133–146. Cambridge University Press, New York.
- Millikan, R. G. (2013b). Reply to Rescorla. In Ryder, D., Kingsbury, J., and Williford, K., editors, *Millikan and Her Critics*, pages 103–106. John Wiley & Sons.
- Millikan, R. G. (2017). *Beyond Concepts: Unconcepts, Language, and Natural Information*. OUP, Oxford.
- Noel, A., Fang, Y., Yang, N., Makrakis, D., and Eckford, A. W. (2017). Using Game Theory for Real-Time Behavioral Dynamics in Microscopic Populations with Noisy Signaling.
- Owren, M. J., Rendall, D., and Ryan, M. J. (2010). Redefining animal signaling: Influence versus information in communication. *Biology & Philosophy*, 25(5):755–780.
- Pfeifer, J. (2006). The Use of Information Theory in Biology: Lessons from Social Insects. *Biological Theory*, 1(3):317–330.
- Piccinini, G. and Scarantino, A. (2011). Information processing, computation, and cognition. *Journal of Biological Physics*, 37(1):1–38.
- Price, H. (2008). Two readings of representationalism.

- Rathkopf, C. (2017). Neural Information and the Problem of Objectivity. *Biology & Philosophy*, 32(3):321–336.
- Rescorla, M. (2013). Millikan on Honeybee Navigation and Communication. In Ryder, D., Kingsbury, J., and Williford, K., editors, *Millikan and Her Critics*, pages 87–102. John Wiley & Sons.
- Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *The American Psychologist*, 43(3):151–160.
- Rouder, J. N., Morey, R. D., Cowan, N., and Pealtz, M. (2004). Learning in a unidimensional absolute identification task. *Psychonomic Bulletin & Review*, 11(5):938–944.
- Sarkar, S. (2013). Information in animal communication: When and why does it matter? In Stegmann, U. E., editor, *Animal Communication Theory*, pages 189–205. Cambridge University Press, New York.
- Scarantino, A. (2015). Information as a Probabilistic Difference Maker. *Australasian Journal of Philosophy*, 93(3):419–443.
- Shannon, C. E. (1948). A Mathematical Theory of Communication (Part 1). *Bell System Technical Journal*, 27(3):379–423.
- Shannon, C. E. (1959). Coding theorems for a discrete source with a fidelity criterion. In *Collected Papers*, pages 325–350. Wiley-IEEE Press.
- Shannon, C. E. and Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press, Urbana.
- Shea, N., Godfrey-Smith, P., and Cao, R. (2017). Content in Simple Signalling Systems. *The British Journal for the Philosophy of Science*, pages 1–27.
- Sims, C. R. (2016). Rate–distortion theory and human perception. *Cognition*, 152:181–198.
- Skyrms, B. (2010). *Signals: Evolution, Learning, and Information*. Oxford University Press, Oxford.
- Stegmann, U. E., editor (2013). *Animal Communication Theory: Information and Influence*. Cambridge University Press, New York.
- Whyte, J. T. (1990). Success Semantics. *Analysis*, 50(3):149–157.