

[Forthcoming in R. Joyce (ed.) *Routledge Handbook of Evolution and Philosophy*]

## **Evolution and information: An overview**

**Ulrich E. Stegmann**

Many biological processes appear to involve the storing, transmission, and processing of information. Heredity is often conceptualized as the transfer of information from parents to offspring and development as the expression of that information. Information also figures at higher levels of the biological hierarchy, e.g., in animal communication and individual decision-making. Evolutionary biology, too, has seen its share of “informational thinking.”

Yet, information is a notoriously slippery concept that often generates confusion and cross-talk. When reading a newspaper we hope to acquire information about an event of interest, in which case information is a piece of knowledge or, perhaps, a proposition. In information theory, a branch of mathematics that can be traced back to Claude Shannon’s (1948) theory of communication, information is concerned with a set of quantifiable, probabilistic features. And then there is the sort of information that one thing can carry about another, e.g., the information that the growth rings of trees carry about the age of trees. This kind of information is known as “natural information.” Traditionally, natural information is construed as a mind-independent feature of the world, something that organisms can exploit in order to guide their behavior. It is therefore also likely to be important for evolutionary processes.

Natural information has received considerable philosophical attention since the 1970s. Fred Dretske’s (1981) book *Knowledge and the Flow of Information* was an early landmark, which shaped much subsequent thinking in the field. The present chapter surveys some of this work, delineating what has been achieved and identifying some of the recurrent challenges. A specific exploration of the evolutionary significance of natural information, however, is a matter for future research and hence beyond the scope of this chapter.

### **1. Evolution and Information**

Before surveying the philosophical work on natural information it is useful to mention some of the points of contact between evolution and information more generally. One of the earliest connections was made by George Williams (1966), who combined natural selection with information in order to define genes. Genes, for Williams, are units of hereditary information that

are subject to a certain amount of selection, where “information” is intended to mean something quite distinct from matter and energy. Williams’ gene concept was criticized for implying a mysterious domain of information (Godfrey-Smith & Sterelny 2008) as well as for privileging genes as the only object of evolution (Griesemer 2005).

Another point of contact was elaborated in John Maynard Smith and Eörs Szathmáry’s (1995) influential book *The Major Transitions in Evolution*. Major transitions are important evolutionary innovations, such as the transition from uni- to multi-cellular organisms. Maynard Smith and Szathmáry argued that the major evolutionary transitions are intimately connected with changes in the way in which information is stored and transmitted. Among these changes is the emergence of non-genetic (epigenetic and cultural) inheritance systems. Another important change is the transition from “limited” to “unlimited” inheritance systems, which essentially differ in the amount of information they can transfer. Maynard Smith and Szathmáry also maintained that the amount of genetic information increases over evolutionary time. The latter claim was later assessed with the help of information theory. In computer simulations, a population of nucleic acids with random base sequences was subjected to mutations and selection for efficient protein binding sites. After 700+ generations, the nucleic acids had converged on non-random base sequences at the binding sites (Schneider 2000). This convergence amounted to an increase in (quantitative) information at the binding sites because the base sequences at these sites had become more and more predictable. The study also undermined a contention popular among defenders of Intelligent Design—namely, that natural processes like mutation and selection are insufficient for increasing information over evolutionary time.

More recently it has been argued that information theory can improve the theoretical basis of evolutionary biology. The theory of natural selection is the formal centerpiece of evolutionary biology, and it is exemplified by Ronald Fisher’s (1930) “fundamental theorem of natural selection.” The theorem describes how the fitness of individuals affects the rate of change in a population’s average fitness. The rate of change is closely related to a population’s growth rate, and it was found that a population’s growth rate can be described with a quantity that has an information theoretic interpretation (Jeffreys divergence). This fact, in turn, allows an information theoretic articulation of the theory of natural selection as a whole. According to Steven Frank (2012), such an articulation is preferable to the statistical formalisms of the classical theory (but see Sarkar 2014; Crawford 2015).

## 2. Introducing natural information

Natural information was introduced above with the example of growth rings of trees. Growth rings are informative in the sense that they can reveal tree age. Dretske (1981) captured this idea as follows:

When a scientist tells us that we can use the pupil of the eye as a source of information about another person's feelings or attitudes, that a thunder signature (sound) contains information about the lightning channel that produced it, that the dance of a honeybee contains information as to the whereabouts of the nectar or that the light from a star carries information about the chemical constitution of that body, the scientist is clearly referring to information as something capable of yielding knowledge. A state of affairs contains information about X to just that extent to which a suitably placed observer could learn something about X by consulting it. (1981: 45)

In short, the basic intuition behind information is that it may yield knowledge. But *how* is it possible that one event can tell us about another? Exactly when is an observer "suitably placed"? Answering these questions is the task of a philosophical *theory* of natural information. In other words, theories of information aim to identify the specific conditions under which one event or state of affairs carries information about another.

Some terminological conventions will be useful in what follows. First, let's distinguish the information-carrying state of affairs (*r*) from the state about which it carries information (*s*). Second, let's refer to the information content, which *r* carries about *s*, with the predicate "... is *F*." Hence, growth rings (*r*) carry information about tree age (*s*), and the information they carry is, say, that the tree is three years old (is *F*), rather than some other age. Similarly, a fingerprint (*r*) carries information about the murderer (*s*), and the information it carries is that the murderer is Moriarty (is *F*), rather than someone else. The task of a philosophical theory of information is then to fill in the dots in the proposition "*r* carries information that *s* is *F* if and only if ...." The dots are placeholders for individually necessary and jointly sufficient conditions.

## 3. Dretske's account

Dretske's (1981) work is a good starting point for surveying philosophical theories of natural information. Dretske argued that *r*'s ability to carry information consists in a probabilistic relation

between  $r$ ,  $s$ 's being  $F$ , and one's background knowledge  $k$ . He articulated the relevant probabilistic relation in terms of conditional probabilities, i.e., the probability of an event *given* that some other event has occurred. Here is Dretske's analysis of (natural) information:

A signal  $r$  carries the information that  $s$  is  $F$  = the conditional probability of  $s$ 's being  $F$ , given  $r$  (and  $k$ ), is 1 (but, given  $k$  alone, less than 1). (1981: 65)

This analysis involves two conditional probabilities. The first is the probability of  $s$ 's being  $F$  given *both* the signal  $r$  and background knowledge  $k$ ; the second is the probability of  $s$ 's being  $F$  given only the background knowledge  $k$ , in the absence of signal  $r$ . Signal  $r$  carries the information that  $s$  is  $F$  just in case the first conditional probability is unity and the second is smaller than unity. Plausibly, both conditions are satisfied in the case of light emitted from distant stars. The probability that a given star ( $s$ ) has a certain chemical composition ( $F$ ) is less than 1 in the absence of knowing what light ( $r$ ) it emits. But the star cannot fail to have that composition if it emits light of a certain quality. Hence, the light carries the information that the star has a certain chemical composition.

Three features of Dretske's (1981) account are worth highlighting. One feature is the reliance on laws of nature. According to Dretske, only states connected by laws of nature, or logical principles, can have conditional probabilities of 1. Stars "must" have a certain chemical composition if they emit light of a certain quality because a star's composition is nomologically related to the emitted light. Accidental coincidences, by contrast, cannot guarantee conditional probabilities of 1. Suppose my neighbor drives to work in the morning shortly before the local bus arrives. Assuming my neighbor drove off just now, there is a high conditional probability that the bus will arrive shortly. But since my neighbor is not connected to the bus by a law of nature, a mere change to the bus schedule diminishes the conditional probability of the bus arriving.

The second feature of Dretske's account is that natural information is always correct. Information is veridical. According to the first condition,  $r$  carries the information that  $s$  is  $F$  only if the occurrence of  $r$  raises to 1 the conditional probability that  $s$  is  $F$ . In other words, the first condition excludes the possibility that  $r$  carries the information that  $s$  is  $F$  in case  $s$  is actually not  $F$ .

The third feature is the role of background knowledge. Dretske illustrates the significance of background knowledge with a "shell game." There are several shells and players must guess which one covers a nut. A player chooses one shell and lifts it to determine whether or not she guessed correctly. If not, she lowers the shell back into place and another player gets his turn. The game

continues until the nut is found. Now suppose there are four shells in the game, and shells 1 and 2 are found empty. At that point a new player joins but is not told about the empty shells. The game resumes with shell 3 being lifted and found empty. At this stage, shell 3's being empty carries different information for the different players. For the original players it indicates that the nut is beneath shell 4, because this is the only remaining shell. However, for the new player it does not carry this information, because his background knowledge does not include the fact that shells 1 and 2 had been turned over already and found empty. So, a given state of affairs can carry information for some observers but not others, and whether or not it does depends on what they know.

All three features proved to be contentious. For instance, the dependence of natural information on a subject's epistemic situation undermines the view that natural information is an observer-independent fact "out there" in the world. It also blocks analyzing doxastic states in terms of natural information, because the appeal to natural information then presupposes a notion of, say, knowledge. Dretske tried to minimize these worries by assuming shared background knowledge among observers, in which case no explicit reference to background knowledge is needed.

Many authors agree that Dretske's probabilistic condition is too demanding (e.g., Loewer 1983; Suppes 1983). Recall that for  $r$  to carry the natural information that  $s$  is  $F$  it must be nomologically impossible for  $r$  to occur if  $s$  is not  $F$ . But many states appear to carry information about others without satisfying this condition. For example, extreme weather events like cold spells can cause trees to form extra growth rings, in addition to the ones reflecting their age. Nevertheless, the growth rings of trees are taken to carry information about their age; an entire discipline (dendrochronology) has emerged on that assumption and proven successful. Furthermore, animals gain information from features that do not guarantee to occurrence of the indicated state (e.g., Godfrey-Smith 1989). Vibrations in a spider's net, for instance, carry information that there is prey without guaranteeing its presence.

Another persistent worry is that there is no suitable notion of probability (Loewer 1983). For example, if probabilities are understood as subjective degrees of belief, then doxastic states cannot be analyzed in terms of natural information, as this would presuppose a notion of belief. Other interpretations also run into difficulties. On a propensity interpretation, probabilities are dispositions of some states to bring about others. However, some states carry the information that others obtain while lacking any disposition to cause them (Demir 2008). For instance, animal alarm signals indicate an approaching predator without in any way causing the predator's approach.

A natural response is to identify and then eliminate the sources of these difficulties. Two such sources are the reliance on background knowledge and probabilities. The counterfactual theory (Cohen and Meskin 2006) therefore seeks to exclude both from a theory of natural information.

#### 4. Counterfactual Theory

Suppose a doorbell is set up in such a way that it only ever rings if someone is at the door. The ringing of such a bell *implies* that someone is at the door. So, we can learn from its ringing that someone is there and, intuitively, the bell's ring carries the information that someone is at the door. According to the counterfactual theory (Cohen and Meskin 2006), the bell's ring carries this information because the ringing is counterfactually related in the right way with people at the door. For if the doorbell only ever rings if someone is at the door, then the following counterfactual is true: if nobody were at the door, then the bell would not ring. The obtaining of this counterfactual relation is seen as constitutive of natural information:

[...]  $x$ 's being  $F$  carries information about  $y$ 's being  $G$  if and only if the counterfactual conditional 'if  $y$  were not  $G$ , then  $x$  would not have been  $F$ ' is non-vacuously true. (2006: 3).

Here,  $x$  designates the information-carrying state,  $y$  the state about which information is being carried, and ' $\dots$  is  $G$ ' denotes the information content. Stipulating that the counterfactual must be non-vacuously true avoids problems resulting from the standard semantics of counterfactuals. Suppose it is metaphysically *necessary* that  $y$  is  $G$ . Then the counterfactual "If  $y$  were not  $G$ , then  $x$  would not have been  $F$ " has an impossible antecedent: by hypothesis, it is necessary that  $y$  is  $G$ , and so there exists no possible world in which  $y$  is not  $G$ . Now, counterfactuals with impossible antecedents are vacuously true, according to the standard semantics of counterfactuals (the Lewis-Stalnaker theories). If vacuously true counterfactuals could ground natural information, then  $x$ 's being  $F$  would carry the information that  $y$  is  $G$ , even though  $y$  could never fail to be  $G$ , independently of whether or not  $x$  is  $F$ .

While Cohen and Meskin (2006) aim to put much distance between their account and Dretske's, they are very similar in one crucial respect. Both imply that natural information is veridical. That is, if  $x$ 's being  $F$  carries the information that  $y$  is  $G$ , then  $y$  is  $G$ . It cannot be the case that  $x$ 's being  $F$  carries the information that  $y$  is  $G$  when in fact  $y$  is not  $G$ .

Since the counterfactual theory does not rely on probabilities, it circumvents the difficulties related to probabilities. And in virtue of identifying natural information with certain counterfactual

relations in the world, information is independent of the epistemic situation of observers. For this reason, natural information can be employed, as Dretske intended, to analyze notions in epistemology and philosophy of mind, such as knowledge and perceptual content.

However, the independence of information from an observer's background knowledge has its downside. Recall the stage in Dretske's shell game when the original players have witnessed shells 1 and 2 being empty and a new player, who does not know about the empty shells, joins the game. The game resumes with the lifting of shell 3, which also turns out to be empty. Intuitively, shell 3's being empty carries the information that the nut is under shell 4 for the original players, but not for the new player, and the difference is due to their diverging background knowledge regarding shells 1 and 2. Dretske's account reproduces this intuition, but the counterfactual theory does not (Scarantino 2008). In order for shell 3's being empty to carry the information that the nut is under shell 4, the following must hold: if the nut were not under shell 4, then shell 3 would not be empty. But this counterfactual is false, because if the nut weren't hidden under shell 4, then it may be hidden under any other shell, including 1 and 2. Put differently, among the possible worlds in which the nut is not under shell 4, worlds in which it is hidden under shell 1 or 2 may be just as close or closer than worlds in which it is hidden under shell 3. Consequently, on the counterfactual theory, shell 3's being empty does not carry the information that the nut is under shell 4, not even for the original players.

The counterfactual theory gives up on Dretske's idea that natural information is a probabilistic relation. But it retains his view that if a state carries natural information about another, then the latter always obtains. Probabilistic theories of natural information take the opposite route.

## 5. Probabilistic information

Dretske's (1981) theory was criticized for being too demanding in requiring that natural information renders a state *certain*. Probabilistic accounts maintain that even imperfectly related events carry natural information under certain conditions, and they then aim to identify these conditions.

Two probabilistic relations should be distinguished from the outset, i.e., degrees of coincidence and probability-changing. The degree of coincidence between one event or state of affairs (*A*) and another (*B*) is the degree to which *As* co-occur with *Bs*. If 20% of *As* co-occur with *Bs*, then *A*'s degree of coincidence with *B* is 20%. For example, if 20% of alarm calls co-occur with predators, then the correlation between alarm calls (*A*) and predators (*B*) is 20%. Put in probabilistic terms, the conditional probability of an approaching predator given an alarm call is  $p(B|A) = 0.2$ .

Alternatively,  $A$  can correlate with  $B$  in the sense of “changing”  $B$ ’s probability. Suppose again that the conditional probability of an approaching predator is  $p(B/A) = 0.2$ . In addition, suppose that predators *always* elicit alarm calls, so that the conditional probability of an approaching predator without an alarm call is  $p(B/\neg A) = 0$ . In this case, the call increases the predator’s probability from 0% to 20% [ $p(B/A) - p(B/\neg A) = 0.2$ ]. Note that in the examples considered so far, coincidence and probability change had the same value (0.2). But this need not be the case. If, for instance, 20% of predators slip through without eliciting alarm calls [ $p(B/\neg A) = 0.2$ ], then an alarm call does not make a predator more likely than it was before. Despite a positive degree of coincidence, there is no change in the predator’s probability.

Probabilistic accounts of natural information come in several versions (Millikan 2000; Millikan 2004; Shea 2007; Piccinini & Scarantino 2010; Scarantino & Piccinini 2010; Skyrms 2010; Scarantino 2015). The differences mainly concern the type of probabilistic relation that is deemed constitutive of information, e.g., whether it consists in coincidence or probability-changing, and whether or not the probabilistic relation must obtain for non-accidental reasons. The most sophisticated account is Andrea Scarantino’s (2015) “probabilistic difference maker theory.”

Scarantino (2015) seeks to capture the central idea of probabilistic accounts, that even imperfectly correlated events can carry natural information, by appealing to probability-changing:

*Incremental Natural Information (INI):*  $r$ ’s being  $G$  carries *incremental natural information* about  $s$ ’s being  $F$  relative to background data  $d$  if and only if  $p(s \text{ is } F | r \text{ is } G \ \& \ d) \neq p(s \text{ is } F | d)$

According to this account, carrying natural information is a matter of one state’s changing (increasing or decreasing) the probability of another. For example, the occurrence of an alarm call ( $r$ ’s being  $G$ ) carries incremental information about a predator approaching ( $s$ ’s being  $F$ ) if the call makes a predator more (or less) likely than in the absence of an alarm call. On the other hand, if the call makes no difference to the probability of a predator, then the call does not carry information about it. Note that the change in probabilities is relative to background data. Background data are sets of propositions which an observer brings to bear and which may or may not be true. In the case of alarm calls it includes the listeners’ assumption that alarm calls make an approaching predators more likely. In the shell game, the background data include knowing that shells 1 and 2 are empty.

Relativizing probabilities on background data turns natural information into a receiver-dependent, 3-term relation. It ceases to be something “out there” in the world that organisms simply encounter. Nevertheless, once the signal and background data are fixed, changes to probabilities ensue independently of an observer’s hopes or opinions. Furthermore, relativizing

probabilities affords a response to the reference problem. Correlations exist only relative to some reference class, and it has been argued that determining reference classes is always arbitrary (Millikan 2013). But relativizing probabilities systematically on receivers' background data allows a non-arbitrary delineation of the reference class.

Increasing the probability of an event does not necessarily make it *probable* to occur (similarly, for decreasing an event's probability). In the example above, an alarm call increases the probability of an approaching predator from 0% to 20%. This increase to 20% does not make it probable that a predator is approaching; that would require the probability to be higher than chance (e.g., 70%). The increase to 20% just means that an approaching predator is now more likely than it was before (without the alarm call). Nevertheless, Scarantino (2015) argues that the overall conditional probability of an event is an important aspect of natural information and has been neglected by other probabilistic accounts. This feature is captured as follows:

*Degree of Overall Support (DOS)*: the degree of overall support provided by a signal  $r$ 's being  $G$  carrying incremental natural information about  $s$ 's being  $F$  relative to background data  $d$  is equal to  $p(s \text{ is } F | r \text{ is } G \ \& \ d)$

In the example above,  $r$ 's being  $G$  raises the probability of  $s$ 's being  $F$  to 70%; its overall support for  $s$ 's being  $F$  is thus  $p(s \text{ is } F | r \text{ is } G \ \& \ d) = 0.7$ .

So far we have looked at the conditions under which one thing carries natural information about another. But what is the content of that information? What does  $r$ 's being  $G$  "tell" us? Any given state of affairs changes the probabilities of many states simultaneously, not merely of one other state. For example, an alarm call increases the probability of an approaching predator, but also of the predator having been detected, the receivers taking evasive action, and so on. At the same time, the call decreases the probability of the predator being successful, of the caller continuing to forage, and so on. Furthermore, the call provides some degree of overall support for all these states. According to Scarantino, they are all part of the alarm call's information content. The formal definition of information content is too complex to reproduce here; suffice it to say that it includes three features: the identity of the states whose probabilities are changed, the amount of change, and the probabilities after the change.

One attraction of probabilistic accounts is that natural information can be shown to do genuine explanatory work (Stegmann 2015). Suppose an organism responds with a type of behavior  $B$  to events of type  $A$ . Since organisms exhibit  $B$  specifically in response to experiencing  $A$ ,  $B$  is part of a dispositional property of these organisms. Suppose we wish to explain, in the first instance, why

the organism responds to a particular event *a* with a manifestation of behavior *b*. The explanation will proceed as follows:

[E 1]

1. Organism *o* has the disposition to respond to *A* with *B*
2. There was an *A*-token (*a*)

Therefore, *o* responded to *a* with behavior *b*

In other words, perceiving *a* together with a disposition to respond to *A*-tokens explains why the organism responded with *b*. This is an ordinary causal explanation and does not appeal to the *A*-token raising the probability of some other event. It therefore does not seem to involve natural (probabilistic) information. However, the role of information becomes apparent once we push our inquiry a step further, specifically by asking why the organism has the disposition in the first place. Let's assume that the disposition is acquired through associative learning. Acquiring the disposition can then be explained as follows:

[E 2]

1. In the past, the conditioned stimulus *A* increased the probability of the unconditioned stimulus *C*.
2. Other conditions necessary for associative learning are satisfied (e.g., surprise and belongingness)

Therefore, the organisms acquired the disposition to respond to *A* with *B*.

Since stimulus *A* increases the probability of *C* (premise 1), *A* carries natural information about *C*'s occurrence. And this fact partly explains why the organism acquired the disposition, as per [E 2]. Natural information is therefore explanatory. Ultimately, natural information is also explanatory of an organism's token behavior *b*, because having the disposition partly explains *b*, as per [E1]. Probabilistic theories therefore show how natural information can play a well-delineated and significant explanatory role.

Of course, probabilistic accounts have their difficulties. Most obviously, the lack of an adequate interpretation of probability, which plagued Drestke's theory, resurfaces here. It remains to be seen whether appeals to objective interpretations of probability (e.g., objective Bayesianism) can fill this gap.

Proponents of probabilistic accounts often believe that they successfully capture the roles of information in the brain and behavioral sciences. Scarantino (2015), for instance, maintains that the information carried by vervet monkey alarm calls is probabilistic information about predators. While alarm calls do carry such information, it is not obvious that this is the kind of information which scientists actually attribute to alarm calls and which figures in their explanations and predictions (Stegmann 2013). This can be seen, for instance, by comparing the information content attributed by scientists with the content predicted by probabilistic accounts. The information content scientists attribute to alarm calls can be fairly narrow, e.g., a vervet's snake alarm call is supposed to signal the presence of a dangerous type of snake, mostly pythons (Seyfarth & Cheney 2003). But on probabilistic accounts, including Scarantino's, the information carried by alarm calls is about much more than approaching predators (see above).

Another challenge concerns "wild tokens." When  $r$ 's being  $G$  correlates with  $s$ 's being  $F$ , then it is possible that some  $rs$  are  $G$  without there being  $ss$  that are  $F$ . Do such "wild"  $r$ -tokens carry the natural information that  $s$  is  $F$ ? If they do, then natural information can be false. This implication is problematic, because it undermines the entrenched view that natural information is always true and, in addition, it requires an alternative criterion for distinguishing between natural information and representational content. On the other hand, if wild tokens are not informative, then carrying natural information cannot be merely a matter of being an instance of a probabilistically related type. And this outcome contradicts the main tenet of probabilistic accounts (Stegmann 2015).

## **6. Natural information as causation**

This section considers a third set of theories of natural information, which analyze information in terms of causation. Causation is usually understood as a relation between things in the world, which either does or does not hold between any particular pair of things. The causal relation that obtains between a token cause and its token effect is known as singular causation.

Singular causation is the essence of natural information, according to Karen Neander's (2013) account of "singular causal information." Where  $r$  and  $e$  are particulars, " $r$  carries the natural indicative information that  $e$  if  $e$  is a cause of  $r$ ." Thus, any given effect carries the information that its cause has occurred. And it does so simply in virtue of the fact that it has been caused by that particular. If a red tomato causes a certain activation pattern (RED) in a sensory neuron, then RED carries the information that there is a red thing. Neander maintains, in addition, that singular causal information can be "forward-looking," i.e., token causes can carry (prescriptive) information about

their token effects. This ensures, for instance, that “motor instructions carry information about the movements they cause” (Neander 2013: 27).

An attraction of singular causal information is its ability to rule out misinformation. By definition, a RED token carries the information that there is something red only if it was actually caused by something red. Furthermore, Neander (2013) maintains that our intuitions about natural information match her account. Recall that extreme weather events may cause trees to form additional growth rings, e.g., a severe cold spell may cause a five year old tree to add a sixth ring. Neander argues that our intuitive judgement in this case would be that the sixth ring does not mean or indicate that the tree is six years old. This intuition would be in line with her account, because the tree’s age does not cause the sixth ring and therefore does not carry information about age.

A potential worry is that biologists and cognitive scientists attribute information to many fewer entities than this account does. For instance, the light-dependent reduction in a light receptor’s firing rate is taken to signal the presence of light, whereas the detachment of part of the retina is not regarded as indicating the presence of the tiny breaks that caused it (except for an ophthalmologist). Yet both events carry singular causal information. The worry is not decisive, however. Probabilistic accounts face the same worry and, furthermore, the cognitive scientists’ more selective use of “information” may be due to pragmatic considerations that are layered on top of a causal notion of information, as Neander suggests.

Another notion built on a causal analysis of information is “mechanistic information” (Bogen & Machamer 2010). Consider the chain of events in mechanisms like protein synthesis. The components occurring at one stage of protein synthesis exert a causal influence on its subsequent stages. The central idea is to identify information with this causal influence, subject to two conditions. One condition is that the components serve an organism’s usual needs, by advancing the mechanism to its end state. The second condition is that the components have long “reach” (roughly, their causal influence extends far downstream and consists in strongly reducing the number of possible downstream consequences). In Jim Bogen and Peter Machamer’s words: “Mechanistic information is the causal influence that entities and activities at one step in the operation of a mechanism exert to select teleologically appropriate results for production in one or more subsequent steps” (2010: 16).<sup>1</sup>

The account draws interesting distinctions between types of causal influence. And, in contrast to Neander’s theory, it appears to construe information as a matter of degree because features like “reach” come in degrees. A number of key issues remain open, however. Does the intended content of mechanistic information concern upstream causes, downstream consequences, or both? Is mechanistic information veridical? There are also questions about the theory’s intended scope.

While it is not promoted as a general theory of natural information, it is meant to account for phenomena that usually qualify as natural information, e.g., the information that the activation pattern of a worm's sensory neurons carries about pressure applied to its skin. Perhaps the theory's scope is more specific, an account of biological information or even just of informational mechanisms.

Ultimately, causal theories of natural information face a common problem. Natural information is often carried by states that are not causes of one another. The fall in a barometer's mercury column indicates a storm, but neither causes the other. Instead they are the effects of a common cause, the drop in atmospheric pressure. Causal accounts rule out such stock examples and, consequently, the possibility of animals extracting information from the myriads of non-causal correlations they encounter.<sup>2</sup>

## References

- Bogen, J. & Machamer, P. 2010. "Mechanistic information and causal continuity." In P. McKay Illari, F. Russo & J. Williamson (eds.), *Causality in the Sciences* (Oxford University Press) 845-864.
- Cohen, J. & Meskin, A. 2006. "An objective counterfactual theory of information." *Australasian Journal of Philosophy* 84: 333-52.
- Crawford, D. 2015. "Sarkar on Frank." *Philosophy of Science* 82: 122-8.
- Demir, H. 2008. "Counterfactuals vs. conditional probabilities: a critical examination of the counterfactual theory of information." *Australasian Journal of Philosophy* 86: 45-60.
- Dretske, F. 1981. *Knowledge and the Flow of Information* (MIT Press).
- Fisher, R. 1930. *The Genetical Theory of Natural Selection* (Clarendon Press).
- Frank, S. 2012. "Natural selection: How to read the fundamental equations of evolutionary change in terms of information theory." *Journal of Evolutionary Biology* 25: 2377-96.
- Godfrey-Smith, P. 1989. "Misinformation." *Canadian Journal of Philosophy* 19: 533-50.
- Godfrey-Smith, P. & Sterelny, K. 2008. "Biological information." In E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, URL = <<http://plato.stanford.edu/entries/information-biological/>>.
- Griesemer, J. 2005. "The informational gene and the substantial body: On the generalization of evolutionary theory by abstraction." In M. Jones & N. Cartwright (eds.), *Idealization XII:*

- Correcting the Model, Idealization and Abstraction in the Sciences* (Amsterdam: Rodopi) 59-115.
- Griffiths, P. E. & Stotz, K. 2013. *Genetics and Philosophy: An Introduction* (Cambridge University Press).
- Loewer, B. 1983. "Information and belief." *Behavioral and Brain Sciences* 6: 75-6.
- Maynard Smith, J. & Szathmáry, E. 1995. *The Major Transitions in Evolution* (Oxford University Press).
- McKay Illari, P. 2011. "Why theories of causality need production: An information transmission account." *Philosophy & Technology* 24: 95-114.
- Millikan, R. 2000. *On Clear and Confused Ideas: An Essay about Substance Concepts* (Cambridge University Press).
- Millikan, R. 2004. *The Varieties of Meaning* (MIT Press).
- Millikan, R. 2013. "Natural information, intentional signs and animal communication." In U. Stegmann (ed.), *Animal Communication Theory: Information and Influence* (Cambridge University Press) 133-146.
- Neander, K. 2013. "Toward an informational teleosemantics." In D. Ryder, J. Kingsbury & K. Williford (eds.), *Millikan and her Critics* (Blackwell) 21-36.
- Piccinini, G. & Scarantino, A. 2010. "Computation vs. information processing: Why their difference matters to cognitive science." *Studies in History and Philosophy of Science* 41: 237-246.
- Sarkar, S. 2014. "Does 'information' provide a compelling framework for a theory of natural selection? Grounds for caution." *Philosophy of Science* 81: 22-30.
- Scarantino, A. 2008. "Shell games, information, and counterfactuals." *Australasian Journal of Philosophy* 86: 629-34.
- Scarantino, A. 2015 "Information as a probabilistic difference maker." *Australasian Journal of Philosophy* 93: 419-443.
- Scarantino, A. & Piccinini, G. 2010. "Information without truth." *Metaphilosophy* 41: 313-30.
- Schneider, T. 2000. "Evolution of biological information." *Nucleic Acids Research* 28: 2794-9.
- Seyfarth, R. & Cheney, D. 2003. "Signalers and receivers in animal communication." *Annual Review of Psychology* 54: 145-73.
- Shannon, C. 1948. "A mathematical theory of communication." *Bell System Technical Journal* 27 379-423,623-656.
- Shea, N. 2007. "Consumers need information: Supplementing teleosemantics with an input condition." *Philosophy and Phenomenological Research* 75: 404-35.

- Skyrms, B. 2010. *Signals: Evolution, Learning, and Information* (Oxford University Press).
- Stegmann, U. 2013. "A primer on information and influence in animal communication." In U. Stegmann (ed.), *Animal Communication Theory: Information and Influence* (Cambridge University Press) 1-39.
- Stegmann, U. 2015. "Prospects for probabilistic theories of natural information." *Erkenntnis* 80: 869-93.
- Suppes, P. 1983. "Probability and information." *Behavioral and Brain Sciences* 6: 81.
- Williams, G. 1966. *Adaptation and Natural Selection* (Princeton University Press).

---

<sup>1</sup> Griffiths & Stotz (2013) also provide a causal account of information, specifically of genetic information. However they place different constraints on the kind of causal relations that qualify as carrying information. Note that the converse project, explaining causation in terms of information, has also been attempted (e.g., McKay Illari 2011).

<sup>2</sup> I thank Gerry Hough, Federico Luzzi, and Stephan Torre for their helpful comments on the draft.