Natural Information, Factivity and Nomicity

Ben Baker

Abstract

Biological and cognitive sciences rely heavily on the idea of information transmitted between natural events or processes. This paper critically assesses some current philosophical views of natural information and defends a view of natural information as Nomic and Factive. Dretske (1981) offered a Factive view of information, and recent work on the topic has tended to reject this aspect of his view in favor of a non-Factive, probabilistic approach. This paper argues that the reasoning behind this move to non-Factivity is flawed and mixes up different problems with Dretske's original view. I show that one of these problems – strictness – has to do with Nomicity rather than Factivity. The other problem – reference class ambiguity – is not solvable just in terms of a theory of natural information. On the Dretske-inspired view I defend, natural information is Factive and Nomic but is insufficient to determine the cognitive or biological content of a natural process. In sum I present an examination what natural information is and what role it can (and cannot) play in our understanding of living and thinking things.

1 Introduction

Natural information plays a crucial role in our scientific understanding of many of the most interesting and puzzling phenomena we find. The notion of information being carried by, and transmitted among states or processes

in the observable, natural world is essential in standard approaches to the study of genetics, nervous systems, and complex organismal behavior. It is largely in terms of the information carried by animal behaviors, strands of DNA, and neural activity that a scientific picture of biological and cognitive phenomena gets filled in. Given that natural information is supposed to play such a crucial role, and given the goal-directed or intentional character of many of the things it is supposed to help analyze, it is a difficult task to spell out precisely what natural information is.

Current accounts of natural information typically take a critical stance with respect to a view offered by Dretske (1981). Dretske's work did much to set the present stage for theories of how the content of perceptions and beliefs states might derive from flows of information. Dretske's view has two notable shortcomings, one being that it is *too strict* and the other being that it fails to provide a way of determining the *reference classes* of information-carrying states. The charge that it is too strict says Dretske's view implies that a many mundane but important cases of information transmission are not cases of information transmission at all. The reference class problem arises from the fact that the information in a signal can only be specified with respect to a determinate range of possibilities and probabilities, and Dretske's view cannot settle what the relevant ranges are in a given case. However, I will argue that the prevailing critical response has misjudged the source of these problems, and that the widespread view of information that results from this misjudgment is confused and lacks the explanatory power that it claims to offer.

A basic aspect of Dretske's view that current approaches typically reject is that it holds natural information to be *Factive.* The Factivity of information means that information pertains to actually existing states or properties, or in other words information indicates what, in fact, happens. Recent analysis of natural information adopt a non-Factive approach on which, broadly, information pertains to probability distributions among possible states or properties. The rejection of Factivity is alleged to help address the problems with Dretske's account [20, 31, 33]. However, the reasoning behind these approaches conflates the strictness and reference class problems. I argue that the move to non-Factivity does not make any progress on the reference class problem and is irrelevant to the strictness problem. The strictness problem is rooted not in Factivity but in Dretske's overly simple conception of Laws

of Nature. This problem can be remedied by modifying Dretske's view to incorporate a more nuanced notion of Nomicity (roughly speaking, physical necessity) that is sensitive to the way initial conditions constrain invariant relationships in nature. The reference class problem is not to be solved by any view of natural information, because it is has not to do with what natural information is, but rather it has to do with what makes certain information explanatorily relevant. Fully appreciating the reference class problem means appreciating where information theory falls short of determining the semantic content that figures in our understanding of biological and cognitive beings. Non-Factive theories of natural information actually exacerbate the reference class problem by massively expanding the space of possible reference classes. In practice, Non-Factive theories obscure the reference class problem, which encourages information-based science that is out of touch with the goal-directed and intentional phenomena of interest. The Dretske-inspired view of information I defend here presents an escape from this confusion and, if adequately supplemented by theoretical resources outside of information theory, can underpin our understanding of the natural occurrence of entities that aim, pursue, avoid, plan, imagine, and more - that is, the natural occurrence of intentionality

I proceed as follows: in section 2, I offer a condensed version of Dretske's view of natural information, and then briefly review several accounts that, in contrast to Dretske, hold information to be non-Factive. I highlight the way that the reasoning offered in favor of the non-Factive views relies on the fact that information-using systems can make mistakes. In section 3, I examine the idea of Law-like or Nomic regularities in nature. I articulate a view of Nomicity that reveals the importance of specific initial conditions associated with any Nomic regularity. This addresses the strictness problem with Dretske's view. In section 4, I argue that the critical reaction to a Factive view of information is based on flawed reasoning; in essence, it tries explicitly to solve the strictness problem, but using unnecessary means, and it ignores the reference class problem, which is actually the deeper puzzle highlighted by consideration of mistake-cases. In section 5, I offer a definition of natural information as Factive and Nomic, supported by the preceding discussion. I then clarify how uncertainty and probability figure in natural information understood in this way. Finally, I show how this Factive approach can better serve the scientific investigation of goal-oriented and intentional systems.

2 Dretske's Theory and the Non-Factive Response

2.1 Factivity and Lawfulness

Dretske aimed to provide for an understand of how perception and knowledge arise from information flowing between the environment and the brain [9]. The account of this information flow expanded on the notion of information developed by Claude Shannon, Norbert Weiner and others [32]. The Shannon model provides mathematical means for calculating stochastic entropy, which has deep ties to the considerably older, thermodynamic notion of entropy. The entropy of Shannon information is usefully glossed as a measure of "uncertainty reduction" [3,21]. Consider the sense in which the result of a coin-flip measurably reduces uncertainty in a lesser degree than the result of tossing a six-sided die - the former reduces two possibilities to one, while the latter reduces six possibilities to one. This is the sense in which the Shannon model says the former result is less informative than the latter. Appreciating the mathematical details of the Shannon model is not important for this discussion. The philosophical questions to do with understanding information as *natural* are beyond the scope of the Shannon model. Since Shannon information is purely a formal measure, it does not say anything to distinguish between a kind of information that is to be found in the nature as opposed to information that is represented in thought or speech. Dretske's concern was to account for information as a natural phenomenon such that it would allow him to naturalistically account for the representational nature of thought.[1]

Dretske's view of information-carrying signals involves two important conditions, **Factivity** and **Lawfulness**. I will focus on Factivity for now, and discuss Lawfulness later. Dretske partly expressed Factivity by reference to Grice's notion of "natural meaning" [15], and by a deliberate use of the word "indicate." In these terms, smoke is an informational signal of fire insofar as it indicates a fire. By contrast, linguistic utterances and conventional signs are, in general, non-Factive. For example, the utterance "there is a fire" and a

---

[1]Going forward I will usually refer simply to "information" without the "natural" qualifier, trusting that my meaning will be clear from context

ringing fire-alarm may *not* indicate a fire - the speaker of the utterance could be mistaken or lying, and the alarm could have been triggered by a prankster or an electrical malfunction. In such misleading cases, Dretske's view is that the utterance and alarm do not carry information about a fire. In his words, "*false* information and *mis*-information are not kinds of information - any more than decoy ducks and rubber rucks are kinds of ducks" (1981, 45). Of course, we sometimes speak of false information or mis-information, but this involves using a non-natural sense of "information," according to Dretske.

The contrast between the Factivity of information and the non-Factivity of certain mental states, especially beliefs, is what sets up *the* central philosophical challenge that Dretske takes on in his 1981 book and several later works. He aims to account for how Factive, informational states of the brain, appropriately related to the world, could help explain non-Factive states like beliefs. In this context, "states like beliefs" broadly amounts to representations that are characterized by a norm of accuracy or truth and that can fail to meet this norm. In such degenerate cases, like false beliefs, the representation is a misrepresentation [10]. In other words, these representations have content that is "trying to get the world right," and which may actually get it wrong, while informational states are never "wrong" about the part of the world about which they are informative.[2].

Slightly rewording the phrasing on 65 and 76-77 of Dretske (1981), I offer the following statement of Dretske's Theory (DT) of natural information:

> **DT:** A signal, b's being in state N, carries the information that s is in state F, if and only if (i) the conditional probability that s is in state F, given that b is in state N, is equal to 1; (ii) this conditional probability relation is fixed by some Law(s) of Nature.

Elaborating further, some part or process in nature, s, is a *source* of information, which can manifest various states, among which one is F. For example, s might be a region of space-time that can either exhibit a fire, F, or a non-fire, G. The event of b's being N is also a source of informa-

---

[2]This is one way of roughly stating "the problem of Intentionality," which has concerned quite a few philosophers besides Dretske [19]
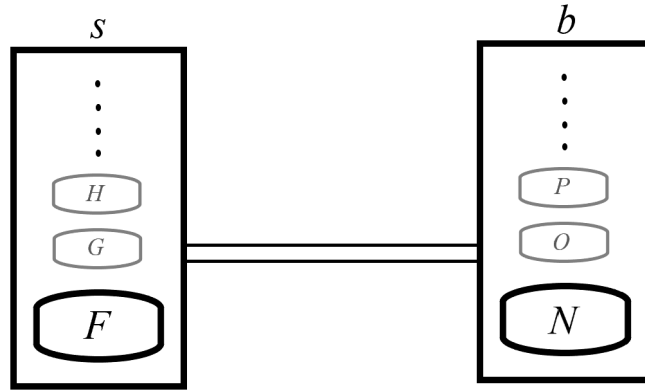
5

Figure 1: b's being in state N carries information about s's being in state F. Smaller, grey boxes and vertical ellipses represent possible counterfactual manifestations of s and b. The double line between s and b represents the non-accidental relation between them; according to DT, it represents a constraint of some Law(s) of Nature.

tion (since b could have been other than N), but here we are not interested in that event as a source but rather as a signal of some other source. For example, b might be a region of space-time that can manifest smoke, N, or not-smoke, O. More relevantly, b might be part of a brain that can exhibit various patterns of neural activity.

Factivity is entailed by DT's first condition, in the required probability of 1. In other words, (i) entails that N indicates F. Since a signal is construed as one among a definite set of possible states of a source, it is more precise to say "b being in state N" carries information, and somewhat misleading to say "N" (all by itself) carries information, but I will sometimes use the less precise phrasing below, for brevity.

The second condition of DT expresses Dretske's Lawfulness requirement on information. More generally, (ii) is a version of a *non-accident* condition. That is, Lawfulness is meant to support the idea that some correlations that are the result of pure chance - mere coincidences or accidents - which are not information-carrying. To illustrate, suppose that there is an ordinary coin that, when tossed many times in an ordinary way, happens to land heads on every single toss. Call this coin "s" and substitute "is F" for the predicate

"is tossed." Call the event of this coin landing "b," and substitute "is N" for the predicate, "lands heads." In such a case there is an observable, perfect correlation between the coin being tossed and the coin landing heads. This is an example of an accidental correlation because, insofar as it involves fair coin tosses, the correlation does not reflect the actual "fairness" of the coin. This correlation does not reflect the physical tendency of the coin to exhibit a certain (distribution of) behavior. The "conditional probability" in DT is meant in the sense that one would say the probability of the coin landing heads is .5, conditional on it being tossed in the air. By contrast, on a simple reading of the "Law of Gravity," the probability that two entities in otherwise empty space will approach each other is 1, conditional on them each having some mass. Pointing out intuitive examples of accidental correlations, like streaks of lucky coinflips, does very little the characterize the correlations that *are* based in information flow, so Dretske looks to Laws of Nature for this purpose, which is rendered by (ii).

It is not self-evident what should be counted among the laws of nature, or which conditional probabilities are determined by those Laws, but the appeal to Laws at least seems to say more than an appeal to "non-accidentalness" alone. For instance, one might reasonably suppose that some Law(s) of Nature ensure correlations between fires and smoke plumes, and that no natural law underwrites the fact that a particular coin lands heads every time it is tossed. [3]

To preface a later discussion, it is worth briefly noting one thing Dretske says about why he invokes Laws of Nature. He claims that the important thing laws do (that accidents do not) is *determine the truth of counterfactual statements.* For example, "if there had been a fire, there would have been smoke" is the kind of claim Dretske relies on Laws for. In his words, "[L]aws have a *modal* quality (they tell us what *must* be the case or what *cannot* happen) that is absent from simple statements of exceptionless correlations."

---

[3]Dretske actually uses different phrases to connote non-accidentalness in different places, and some of these are weaker-sounding than (ii). At times he speaks broadly of "nomic dependence" (1981, 75-76). He also relies considerably on an example where informational relations are fixed by conventions among regular partners in a card game (70), which seems not to involve Lawful correlations in the strong sense DT suggests. Still, Dretske does explicitly appeal to Laws at times, and I believe DT reflects how Dretske's approach has generally been received by others in this literature.

(ibid, 77).

Now, with Dretske's theory in hand, let us turn to some dissenting views.

2.2   Garden Variety Correlations

A prevailing criticism of Dretske's (1981) account of information transmission is that it is too strict [11,14,20,25,31,33]. In short, these critics argue that DT rules out many clear cases of information transmission, including exactly the sort of cases that are supposed to be relevant to a naturalistic understanding intentionality and knowledge.

Ruth Millikan (2004) gives clear expression to the reasoning underlying the common critical reaction to DT [26]. She argues that, insofar as people and animals seem to cognitively rely on information signals, these seem *not* to be signals that carry a lawfully guaranteed indication of some state of the world. To illustrate, she considers the example of a rabbit who hears that there is a predator in its immediate environment. This is an important sort of case since all parties here agree that information about the presence of the fox is carried by its body, transmitted through various structures in the air - light patterns, sound waves, airborne chemicals, etc. - and eventually transmitted to some neural activity in the rabbit. This is supposed to help explain how the rabbit has some intentional state about the fox. Millikan has the following to say about such a case:

> "[N]o natural law can require it to *be* a predator that causes [a rabbit's] predator detectors to fire. Whatever information channel she uses, it is always nomically possible that non-predators should exist who would activate it. Suppose for the sake of the argument (though very implausibly) that there are unbreakable natural laws that concern the effects of foxes on rabbit sense organs. Still, there are surely no laws that nothing *else* could possibly produce these same effects on the rabbit sense organs." (2004, p.33)

In Millikan's view, taking N to be any activity of a rabbit's brain you like, it must be possible for N to be caused in the absence of any fox, say, by something that looks, sounds, or smells like a fox. Millikan's basic conclusion is that DT is untenably strict because it cannot substantiate basic cases like that of a rabbit's brain carrying information as to the presence of predators.

Nicholas Shea enlists Millikan's reasoning in arguing that it is "correlational information" that underlies intentional content, as opposed to information as defined by DT [33]. Importantly, the relevant "correlations" can be imperfect ones, so Shea's notion of information is non-Factive. He makes this clear when he says "instances of a type [] R which carries correlational information about C can be tokened even when C does not obtain" (420). (To translate, change "R" to "b is in state N" and change "C" to "s is in state F"). So, according to Shea, the rabbit's brain state, N can carry information about a predator being around even when, in fact, there is no predator around, provided the relevant correlation exists. This is supposed to help explain how it is possible for the rabbit to have an intentional state whose content does not line up with the facts, that is, how the rabbit can misrepresent things.

Shea's view also includes a non-accident condition. He does not refer to "laws of nature" but says that the relevant correlations are fixed by some "common natural reason." This seems to be a weaker requirement than (ii) in DT, but Shea does not offer a detailed account of what a common natural reason is. The main point here is that the correlations Shea's view relies on are not meant to be mere frequencies, as in the earlier case of a streak of coin-flips landing heads. Rather, Shea wants to rely on "objective" probabilities, "like the 50 percent chance that a lump of 4.5 billion atoms of uranium-238 will emit an alpha particle in a year" (ibid).

Even assuming the probabilities associated with atomic radioactivity are universal, it is much less clear how to make sense of more relevant cases. How does one assess the "objective probability" that some non-predator causes pattern N in a rabbit's brain? Is the probability determined by *all* the non-predators in the universe capable of making it the case that b is in state N, under all possible circumstances? That is implausible, since there are an incomprehensibly large number of non-predators with that capacity, so such a calculation would not yield the information that we antecedently

suppose the rabbit relies on. Anyway most of those non-predators seem not to commonly or naturally interact with rabbit brains, which suggests Shea means to exclude them from consideration. However, if we are to relativize to some non-universal domain in order to determine the relevant probability - say, considering only events on Earth, or in the rabbit's local habitat, or in the habitat of its progenitors - this immediately raises the question of why *that* relativization fixes the "objective" probability relevant to the rabbit's intentional state.

This sort of worry about the appropriate relativization or *reference class* is notorious, and will feature in this discussion in several places [17, 18]. In section 4 I try to clarify the significance of this problem for a view of natural information.

2.3 Probabilistic Information

Andrea Scarantino and Gualtiero Piccinini offer the most sustained and direct argument I have found against Factivity [31], and Scarantino (2015) presents a detailed theory of information as a "probabilistic difference-maker" [30]. Kraemer also rejects condition (i) of DT and defends an account of probabilistic information [20]. Like Shea's, their views are explicitly motivated by observing that information-users sometimes respond to signals that do not indicate whatever the response appears to be a response to. On this score, Scarantino and Piccinini say "[o]rganisms do make mistakes, after all. Some mistakes are due precisely to the reception of probabilistic information about events that fail to obtain" (2010, 319). Probabilistic information content is thus not a fact, but is rather the *extent* to which a possibility is more or less probable. Kraemer also points to the problematic strictness of Dretske's account and proposes that "many natural signs carry information about the *probabilities* of certain occurrences" (2015a, 145). For instance, these philosophers would suggest that the rabbit's brain-state, N, might just carry the information that there is a predator around with some probability, p, whether or not a predator is actually there.

Scarantino and Piccinini's notion of probability also contains a non-accident condition in terms of "reliability;" only reliable correlations count for the pur-

poses of probabilistic information. Also, Scarantino's (2015) use of the term "difference-maker" suggests a requirement on informational relations that shares at least a family resemblance with condition (ii) of DT. Exploring the "difference-maker" idiom in depth is beyond the present scope, but I note that, intuitively, whether there is fire makes a difference to whether there is smoke whereas whether a certain coin has landed heads many times in the past does not make a difference to whether it will land heads on a particular future toss [4]. Scarantino and Piccinini, like Shea, avoid appealing to laws of nature but do not precisely articulate what distinguishes the non-accidental (reliable) correlations as such. Scarantino and Piccinini say that reliable correlations are the sort that "information users can count on to hold in some range of future and counterfactual circumstances," and largely leave it at that (2010, 318).

In Scarantino's (2015) theory, an informational signal implies an incremental change in the probability of an event *relative to a prior* probability, where this prior is fixed by certain background data. So, strictly speaking, a signal like b's being in state N is not said to carry information as to how probable it is that a predator is nearby, but rather information as to how much more or less likely it is that a predator is nearby, compared to some baseline. This will sound a lot like Bayesian confirmation theory to those familiar with it, because it is. Scarantino's view aims to smoothly incorporate the insights of Bayesianism alongside a Shannon-inspired measure of information. Here is his formulation (2015, 423):

> **Incremental Natural Information (INI)**: b's being in state N carries information about s's being in state F, relative to background data d, if and only if $\Pr(s \ is \ F \mid b \ is \ N) \neq \Pr(s \ is \ F \mid D)$

The vital role played by background data in INI illuminates the relativization issue that is obscured in an unadorned appeal to objective probability. According to INI, what information N carries in a rabbit's brain depends directly on what the background data is. Of course, how to identify the relevant background data is non-trivial. Again, I leave deeper examination ofthis issue for section 4.

---

[4]see [36] on difference-making

2.4 Signalling Games

There are deep parallels between the puzzle about the intentional content of cognitive states and one about how to determine the content of the communicative behavior of living things. Animals, plants and even bacteria manifest behaviors that appear to serve as signals to other creatures (especially conspecifics), which might allow those creatures to find food, realize reproductive opportunities, or avoid predators [24, 35]. These behaviors are described in both scientific and informal contexts as signs that communicate contents along the lines of "food is over here," or "I am your potential mate," or "danger from above." What determines this content?

Brian Skyrms' theory of the information content of biological signals uses a game-theoretic framework, building on the work of David Lewis [23, 34]. Lewis was analyzing the meanings associated with actions in a "signalling game." In rudimentary form, such a game involves two players, "Sender" and "Receiver," who each get certain payoffs (or incur costs) that depend on both the state of the world and on their actions. One round of the game goes as follows: only Sender observes the state of the world, then Sender chooses an action, then Receiver observes Sender's choice and must choose an action. Receiver's choice given the state of the world determines the payoffs that Sender and Receiver realize. Lewis' key insight is that sometimes, given the right arrangement of possible actions and payoffs, Sender's rational (payoff-maximizing) strategy is one that allows Receiver to learn something about the state of the world from Sender's choice. That is, it can be to Sender's benefit to correlate her actions with what she observes, allowing Receiver to do better than randomly guess the state of the world. (This requires the typical game-theoretic assumptions that the available options and the rationality of the players are common knowledge among them). Lewis uses this approach to try to account for the meanings of conventional signals, especially linguistic ones. Skyrms draws on this approach to develop a theory of the content of communications in the biological realm generally.

One oft-discussed example of animal communication is that of vervet monkey alarm calls, which differ depending on the type of predator that a monkey sees, and allow monkeys in earshot to flee to bushes, or treetops, or wherever is most appropriate for a specific type of predator. In short, vervets seem

to *warn* each other as to which predators are around. To simplify for the purposes of describing this case in terms of a signalling game as per Skyrms, suppose the range of possible states of the world is "Snake," "Eagle," and "No Predator." Suppose also that Sender monkey, who sees the world state, can take three different actions (calls), and that a Receiver monkey, who hears the call, can also take three different actions. Finally, suppose that each state of the world is associated with a single high-payoff option for Receiver, so there is a clear best choice given each state. Depending on Sender's payoffs, her optimal strategy might involve correlating her different calls with the different states in a way that supports Receiver taking a strategy of correlating his actions with her calls, thereby reliably getting high payoffs.

Skyrms' view is that biological signals like alarm calls carry information insofar as a Receiver's actions warrant attributing to the Receiver an implicit *probability assignment* as to the state of the world, conditioned on the received signal. Skyrms' information content is thus a vector of probabilities over possible states. For example, $[p, q, (1\text{-}p\text{-}q)]$ could be a vector describing the information in an alarm call, where p is the probability assigned to Snake and q is the probability assigned to Eagle, and No Predator is the remainder. Importantly, this model does not require perfect correlation between Sender actions and states of the world, so these vectors involve probabilities less than 1. In other words, Sender can (at least in some relevant cases) be expected to employ a "mixed strategy," where she sometimes takes different actions for the same state of the world. This renders the "message" in her actions somewhat ambiguous as to the state of the world. For instance, Sender's strategy could be to have one call that is highly correlated with snake but that she occasionally uses in the No Predator state, as it could occasionally be in her interest for Receiver to act as though there is a Snake when there is none. This illustrates how information content can fail to indicate the state of the world, on this approach. The non-Factivity of this approach is also plain from the way information content is roughly equated to *Receiver's guess* about the state of the world - the relationship between the received signal and the world is mediated by what *Sender chooses*, having observed the world.

On Skyms' view, information is relativized to a game theoretical model in a way analogous to how, in Scarantino's account, it is relativized to a particular set of background data. This is because the information content

in a biological signal is determined by the details of the game used to model the phenomenon. Which game models are appropriate to which biological phenomena is a separate, difficult question, which looks a lot like the reference class problem. Also, in specifying that the players play payoff-maximizing strategies given common knowledge, one effectively stipulates the relevant modal facts (such as what Sender would have done if she had seen a different state) so the question of whether accidents carry information does not arise on this approach.

I have now surveyed several non-Factive views of information and noted how each relies on a kind of reference class and on the exclusion of merely accidental correlations. I now shift focus to "non-accidentalness" itself, which will let me spell out the view of information I defend.

3 Non-Accidentalness as Nomicity

3.1 What Matters is Modality

Recall that Dretske was motivated to rely on laws of nature as a way of determining the truth of counterfactual claims. Similarly, Scarantino and Piccinini expound on the "reliability" of a correlation by pointing to what would happen in some counterfactual cases, and Shea also suggests that "common natural reasons" underwrite relevant counterfactuals. So there is some consensus here that the important difference between accidental and non-accidental covariance is the counterfactual implications that come with the latter. Everyone here agrees, for instance, that if smoke carries information about fire, the smoke-fire covariance implies the following: if there were no smoke there would be no fire (or at least, fire would be less probable than if there were smoke). By contrast, the fact that a certain coin has landed heads 100 times in a row does *not* imply that the 100th toss would have (probably) landed heads even if it had been tossed somewhat differently.

A reason to think counterfactual implications are required of information-bearing entities is that this makes sense of the potentially *guiding* or *re-*

vealing character of information. This is expressed by Dretske's claim that information is *what one can learn* from a signal, and expressed by Scarantino and Piccinini's claim that reliable correlations can be "counted on," and expressed by the much older idea that objective probabilities must be a "guide to life" [4,5]. Information must be presumed to be bound up with the counterfactual structure of the natural world, or it would not help explain the way information-users accurately aim at things, reliably predict things, correctly infer things, etc. In short, counterfactual implications make information useful.

One might hope to avoid worrying about non-accidentalness by construing information directly in terms of the relevant counterfactuals, avoiding any reference to Laws or physical necessity or probability. Cohen and Meskin (2006) take this approach. Their basic thesis can be stated as follows:

>   **Counterfactual Theory of Information (CTI)**: b's
>   being N carries information about s's being F if and only
>   if the counterfactual "if s were not F, then b would not be
> N" is non-vacuously true. (335)

The qualification "non-vacuously" is meant to block the implication that information is always carried by any signal about any necessary condition. Without this qualification, for instance, CTI would imply that every signal carries the information that 2+2=4.

CTI is a Factive notion of information, in that b's being N and s's being F are construed as actual events (ibid., footnote 2). Note also that anything that carries DT-information carries CTI-information, given the counterfactual implications of laws of nature. However, the converse may not hold. CTI is a weaker thesis than DT in that it does not commit to laws of nature as determining the truth of the relevant counterfactuals. CTI is silent about *what makes* the counterfactuals non-vacuously true. This silence makes CTI unsatisfying in the present context. Because the truth of relevant counterfactual statements is simply asserted, the "guiding" or "revealing" character of information is left as something of a mystery. To ask "In virtue of what do these counterfactuals hold?" is to ask about non-accidentalness, which CTI offers no view of. If not by appeal to laws of nature, some other positive notion of non-accidentalness is needed to say where the counterfactual

implications of information come from.

I will now argue that a careful understanding of physical necessity, i.e., law-like invariance i.e., *Nomicity* is the relevant, positive notion of non-accidentalness. To facilitate my discussion of Nomicity, I want to consider three variants of an example I will call **Thermometer Information:**

A. In a typical thermometer, a volume of mercury, V, linearly covaries at rate r with the temperature, T, of the surrounding space. (V carries information about T)

B. A volume of mercury, V, linearly covaries at rate r with the temperature, $T2$, of a location on a different planet. (V does not carry information about $T2$)

C. In the vacuum of outer space there is a volume of (quite solid) mercury, V, which does not linearly covary with the surrounding temperature, $T3$. (V does not carry information about $T3$)

One can make uncontroversial judgments about the relevant informational relations in these cases, which I stated in parentheses. What we want to know is why they are the right judgments. Case (A) is a standard one of non-accidental covariation. Case (B), by contrast, exemplifies accidental covariation - the relationship is "mere chance." Case (C) involves some mercury and a surrounding temperature - apparently the same kind of entities as in (A) - but the linear covariation is missing because the mercury is frozen. Thus, it appears the non-accidental relationship that fixes information transmission in (A) must not hold in (C). But one might have thought that non-accidental, information-fixing relationships hold everywhere and always. So what gives? Surely a standard thermometer carries information about the temperature, but what determines the conditions of this information-carrying relationship? Note that we cannot even say that the generalization in (A) holds "usually" or "on average" if those are interpreted in basic physical terms, since much more of the universe is below mercury's freezing point than above it. Of course, "usually on Earth" is another matter, but why would non-accidentalness itself, insofar as it determines the content of natural information, be relativized in such a way? Again the reference class

problem rears its head. To get clear about what information is transmitted where, we need a more exact view of non-accidentalness.

3.2   Law-Like Invariance, Nomicity

Laws of nature are supposed to describe regularities of the most general and fundamental kind, so it would be nice if information was determined by such laws. Since information is supposed to inhere in the structure of nature itself, flowing among nature's parts before any intelligent being opts to make use of it, any laws governing the operations of *all* of those parts would be safe ground on which to base a theory of natural information. The problem is that it is extremely implausible that, as Millikan put it, laws of nature concern the effect of foxes on rabbit sense organs.

Even if laws of nature do not concern things as specific as rabbit brains, the idea that laws are *perfectly* general regularities, nowhere and never admitting of exceptions, is too simplistic. Such a simplistic view is suggested by the way philosophers have often described laws as logical schema, especially as universally quantified conditional statements [1, 8, 37]. For example, a classical statement of a law might be "all smoke comes from fire," or "if ambient temperature increases by X, the volume of mercury increases by $X*r$." Thinking of these as expressing laws fits with a view of laws as universally applicable, but such a view is not tenable. As I illustrate below, having a clear view of the significance of *initial conditions* in laws of nature allows for a clearer view of Nomicity as non-accidentalness, which resolves the kind of puzzle brought out by Thermometer Information. For more thorough accounts of laws that clearly lay out their non-universality, see [6, 22, 38], and for analyses specifically of the important pairing of initial conditions with statements of laws, see [2, 7, 12, 39].

The most widely agreed on examples of laws of nature describe how the simplest physical quantities interact, and they are typically expressed in the mathematical terms that make up our most general models of the physical universe. Laws of this sort come with a range of initial conditions - conditions in which we might observe the system being described and under which the law is expected to hold. There are parts of the universe we have measured

17

that exhibit extreme conditions that exceed the bounds assumed by statements of laws (like immediately after the Big Bag), and parts of the universe we have not measured that are theorized to lie outside these bounds as well (like the inside of a Black Hole). General Relativity, for example, is thought not to apply over the smallest measurable distances due to quantum effects, and at very high energy scales the Strong, Weak, and Electromagnetic Forces appear not to behave remotely like our models of them would have it [39]. If a model of some interacting physical quantities needed to accurately describe things in literally every condition in the universe in order to be a law of nature, then we would have no such model.

Laws are typically formulated by scientists as (systems of) differential equations, which differ from universally quantified conditionals in several important ways. These differential equations involve variables that mutually shape the way the variables change over time. Each variable represents some aspect of nature and can take a wide range of values while still supporting the invariant relationship specified by the equation(s). What a law says is invariant is the relationship among the variables specified by the equation(s) - what is invariant is *not* a generalization about concrete, observable entities. Such generalizations, like "smoke comes from fire," can be thought of as being described by certain combinations of certain values of the variables that figure in a law. This means that a lawful invariance can be realized in starkly different ways, which are represented by starkly different solutions to the differential equation(s). It does not make sense to think of all of these possible solutions as "instances" of an (absurdly convoluted) conditional statement, because the solutions range over phenomena that are so qualitatively diverse. Crucially, differential equations typically only *have* solutions when the variables fall within certain ranges; roughly put, if one value is too high or low, there is no guarantee that the rest of the values can be filled in so as to make the equation(s) hold. In sum, the mathematical statements in physics which have most claim to "law" status can be expected *not* to apply to certain situations where some initial condition is not met.

To further illustrate, if there were a law concerning the relationship between the volume of mercury and the surrounding temperature, it might specify that the two variables, V and T, *do* perfectly covary at rate C under a range of initial conditions. These conditions would be exceeded in the vacuum of space, which would explain why the law does not apply in

case (C) of Thermometer Information. However, our broadest theories of the physical universe do not include laws specifically about mercury and surrounding temperature. Rather, the law(s) relevant to that case involve highly general equations with variables whose values can describe both solid and liquid mercury, and countless other measurable kinds of stuff. These laws hold both in typical households and in the vacuums of space, and so the invariance described by the relevant equations is manifest both in the observable covariance in (A) and in the lack thereof in (C). This is somewhat opaque because to represent what is invariant across these two cases requires a degree of physical abstraction and precision that is difficult or impossible to achieve in a spoken language (which is one reason calculus is important).

Most of the regularities that we talk about and that inform how we live our lives are like that in (A): they are *law-like* insofar as they involve invariances over a range of initial conditions, but compared to the invariances commonly dubbed "laws of nature" they are far less general. These prosaic regularities have a much narrower domain of application than Laws and are less theoretically relevant for the empirical study of other very broad generalizations. The difference here is a matter of degree.

Recognizing the non-universality of laws allows one to appreciate the sense in which they are of-a-kind with regularities in all manner of domains other than that of making highly abstract models of physical systems. Covariance between smoke and fire, between foxes and rabbit brains, and between fundamental particles all share the feature of being invariant under a range of initial conditions. The point is that this is all one should mean by "physically necessary" or "Nomic." A law is guaranteed to hold in a very wide range of conditions, but any other pattern of covariation is no less physically necessary because it is only guaranteed in rare circumstances. A striking correlation among the outcomes of successive flips of a coin is rightly regarded as random or accidental insofar as it was not guaranteed under the conditions that it occurred. In other words, non-accidentalness is Nomicity, which characterizes both extremely general regularities, like gravitational attraction, and extremely specific ones, like the effect of a fox on a rabbit's brain.

Recall that condition (ii) of DT required that information flow be "fixed by some law(s) of nature." Among the dissents to DT I reviewed above, Millikan's was particularly critical of (ii), but the others also at least tacitly

rejected it in favor of some weaker-sounding non-accident condition. I agree that we should reject (ii) for the reason Millikan offered; there are not laws concerning many things involved in interesting cases of information transmission, like that between a fox and a rabbit's brain. But there *are* Nomic relationships between foxes and rabbit organs. Requiring that informational relations be Nomic therefore does not imply an unacceptably strict view about what informational relations exist.

I will now return to reconsider Factivity. I argue that, having recognized the problematic strictness of DT is resolved by the replacing the appeal to laws by an appeal to Nomicity, Factivity should not be abandoned.

4 Nomic, Factive Information

4.1 Mistake Cases and Factivity

To briefly reiterate, condition (i) in DT says that if b's being in state N carries information about s's being in state F, then the probability of s's being in state F, conditional on b's being in state N, is 1. Less technically, DT says that b's being N indicates that, in fact, s is F. I outlined several accounts that reject this condition partly on the basis of considering cases involving *mistakes*. That is, these accounts take the observation that information-users make mistakes to count in favor of a non-Factive view of information. However, this reasoning is critically flawed. To illustrate, I consider the following two variants of the case of **Fleeing Rabbit:**

A. A pattern of air-vibration caused by a *stalking fox* reaches a nearby rabbit's ears and causes a brain state N in the rabbit. N then figures in further brain processes that normally generate the rabbit's fleeing behavior, and it does flee.

B. A pattern of air-vibration caused by a *fallen branch* reaches a nearby rabbit's ears and causes a brain-state M in the rabbit. M then figures

in the same processes as in (A), and the rabbit flees. No predator is
nearby.

Let us stipulate that the pattern of air-vibration at the rabbit's ear and
its brain activity are exactly similar in terms of their local physical proper-
ties. So, everything after the fox's step is indistinguishable from everything
after the branch's landing - that is, indistinguishable from the perspective of
a well-positioned observer who has superbly high-resolution devices for dis-
tinguishing patterns of air-vibration and rabbit brain activity. This means
that N and M are constituted by the same rates of neural firing in the same
neurons, which figure in the same larger brain processes in a rabbit with
the same learning history, and the rabbit's sense organs receive the same
impinging pattern of energy.

Allow me to reconstruct the line of reasoning that I take to be the com-
mon objection to DT. To start, we can assume N carries predator-related
information, since we are assuming the information in N supports ascribing
predator-related intentional content to the rabbit. Now add the following
premise; N and M must carry exactly the *same information*. One might
think this premise is supported by the similarity between N and M that I
just stipulated, or simply by observed similarity in rabbit behavior across
(A) and (B). That is, one might think that the sameness of what happens
locally around the rabbit brain implies that N and M must be information-
carriers of the same basic type. If this were right and N and M do have the
same information content, then M must carry predator-related information
even in the absence of a predator, since N does. Thus, the non-Factivity of
information would follow.

However, the stipulated similarities between N and M do *not* entail that
they carry the same information. Physical shape and local causal relations
do not fully specify information content, because they do not exhaust Nomic
relations. The rabbit's brain activity is Nomically related to innumerable
parts of the world the rabbit lives in, ranging widely in temporal and spatial
scale. What happens in the rabbit's brain and nearby its sense organs is an
infinitesimal portion of all of the Nomic regularities that involve the brain
activity described by N and M. By assumption, the two brain states have
different causal relations to moderately distal objects - a fox and a fallen

branch - so there are some things N is Nomically related to which M is not, and vice versa. In short, one could say that the two brain states are Nomically differentiable. This suggests N and M should be differentiable in terms of the information they carry.

Here it may be worth emphasizing the assumed naturalness of information. Recall that the non-Factive alternatives to DT discussed above share with Dretske the idea that it is not *observed* or *perceived* correlations that determine information content - that would leave us with a vicious circularity when it came to using information flows to explain the content of observations or perceptions. Rather, information inheres in the counterfactually robust structure of the physical world itself - in Nomic correlations as opposed to accidental ones. Since Fleeing Rabbit (A) and (B) do not describe Nomically identical situations, the sense in which N and M might be classified as states of the "same type" cannot be just based in their Nomic properties.

Recognizing the way Nomic regularities involve initial conditions, replacing DT's condition (ii) with a Nomicity requirement supports a distinction between information flows across Fleeing Rabbit (A) and (B). We can assume that under *some* (relatively narrow) initial conditions, N *does* guarantee the presence of a fox, and that M is not subsumed under that same Nomic regularity because it occurs outside of those initial conditions. This points to how vast and diverse are the arrays of Nomic regularities involving N and M, and how inadequate local similarity is as a measure of informational similarity. Understanding information to come with all Nomic relations yields an extremely inclusive (not strict) view of information, which both resolves one of DT's main problems and provides the means of distinguishing N and M in terms of the information they carry. This undercuts the reasoning behind the non-Factive approaches above.

Thinking back to Figure 1 (p.5) and supposing we observe b's being N (or M) in the Fleeing Rabbit's case, we can see that the information content is given by the distribution of possibilities at the information source, s's being F. This source can be *any* probabilistically complementary set of events that N (or M) is Nomically connected to. So b's being N presumably carries information about the presence of a fox, about patterns of air movement in the rabbit's ear, about the presence of a working rabbit brain, about the presence of a life-sustaining planet, and much, much more. Each measurable

22

quantity of information can be specified with respect to the initial conditions of one or another Nomic regularity that N stands in. This extremely inclusive view of the Nomic basis of information fits with the pre-theoretical idea that information, understood as a natural, objective commodity, is what a hypothetical observer *could* learn from a signal, where such a hypothetical observer might occupy any perspective we can imagine. So, such an observer might be interested in any s, described in any manner of F-ness, G-ness, etc., that stands in any Nomic relation to b's being N. Therefore when we speak of "*the* information" that some neural activity carries - *all* the information - we are speaking of an incomprehensibly huge and diverse multitude of things, each involving some Nomic regularity given some initial conditions.

Since N and M cannot be classified as states of the "same type" just in terms of their inherent informational properties, what might justify their being classified together? Well, they appear to play the same role in a complex system that generates Rabbit-fleeing behavior. One might therefore think that, despite having different informational properties simpliciter, N and M carry the same information *relevant* to how that behavior is generated. If true, this would support the non-Factive approach on a line of reasoning much like the one I rebutted above. However, note that this requires a shift from the question of "what information is there" to the question of "what information is relevant" (for understanding some phenomenon of interest). The latter question just states the reference class problem in terms of possible sources of information. That is, to classify N and M as relevant-information-carriers of the same type involves selecting probabilistically complementary sets (a reference class) that specifies the kind of transmission or the "channel" we are interested in, and disregards myriad other information flows that occur in Fleeing Rabbit (A) and (B). Thus, classifying N and M together assumes an answer to the reference class problem that makes them relevant in the same way.

If one selects a reference class that roughly reflects the epistemic position and interests of an individual human, one will readily identify informational similarities between N and M. That is, across the various contexts in which we would observe a rabbit brain, N and M would look, to use, like the same kind of state - we could learn the same sorts of things from each. For instance, we could learn from each that the rabbit is about the exhibit predator-avoidance behavior. But referring to what we can tell apart brings

our own reference frame into the picture, so to speak. It smuggles our own intentionality into what is ultimately supposed to help explain how a brain, body, and environment interact to give rise to intentionality. The assumption of a human observer-centric reference class in a theory of natural information is therefore illicit. To *define* information in a way that makes N and M information-carriers of the same kind would be to relativize information to human intentionality and give up any hope of explaining the nature of intentionality in terms of information.

Non-Factivity does not resolve the reference class problem; which background data or natural correlations are relevant is not, in general, a transparent matter. By one measure, non-Factivity makes the reference class problem worse, since one needs to specify both the kinds of entities involved (brain activity, b and certain environmental features, s) and the populations of instances of those entities that fix the probabilistic information content. Of the many things N carries information about, we need some reason to hone in on fox-ness, and also a reason to hone in on the particular distribution of foxes and non-foxes on which the conditional probability of N versus not-N provides the information content. Non-Factivity is therefore not well-motivated as a response to DT's problems.

By my lights the move to non-Factivity loses touch with the agenda of understanding thinking processes naturalistically, adopting an assumption about relevant reference classes that relies on the intentionality of human investigators. This suggests the non-Factive views are better seen as theories of non-Natural information. In many contexts, where a certain investigative perspective can properly be taken for granted, these theories might be quite useful. But in the context of defining information as it figures in the explanation of how intentionality arises in nature, one needs to give due attention to the reference class problem, on pain of vicious circularity.

Stating the reference class problem in terms of the question "what information is relevant...?" should make clear that the solution to this problem is not to be sought in information-theoretic terms. The similarity between (A) and (B) is not just a matter of what information enters the rabbit brain, but is rather something about some of the information there. This "something" is what makes the predator-related information relevant, it tells us which classes of rabbit brain-states and Nomically related environmental events we

are interested in, and it is not something that a good theory of information should capture.

Importantly, the relevance-fixing similarity between (A) and (B) could be a feature of processes that N and M figure in despite carrying different information. In other words, whatever solves the reference class problem could explain why M does something like N even though the latter carries Factive information about a fox and the former does not. There is significant precedent for relying on a notion of *function* for to this end. Later Dretske (1988) and many of his interlocutors rely on a notion of biological functions as selected by evolution or learning [13, 14, 16, 27–29]. Whether and how this idea can be coherently spelled out is beyond scope here. But supposing it is the fact that N and M serve the same function and that this satisfactorily accounts for what is the same across (A) and (B), there is nothing problematic about M carrying different (and also Factive) information from N, as long as M functions in the same way as N despite this informational difference. If it is *how the information is used* by a larger system that settles reference class issues, it is reasonable to think the rabbit uses (Factive) information about a predator in (A), whereas in (B) the rabbit uses M as if it carried that information, though it does not. It could be that the complex of perceptual and motor processes that go into defining the function of N are such that (Factive) information about the movement of the air (or some other proxy for F) is used as if it were information about a predator.

It is not obvious that a theory of function and the implicated notion of selection evolution or learning will ultimately allow one to make sense of intentionality in naturalistic terms, but my focus here is on understanding information, assuming information is to figure in that explanatory project. Since resources from outside of information theory are needed to determine reference classes anyway, adopting a non-Factive view of information is confused as an attempt to deal with cases like (A) and (B).

4.2   Nomic, Factive Information and Probability

My analysis here suggests processes of perception and intentional action are to be partly understood in terms of information flows within a nervous sys-

tem and its environs, acknowledging that it is a substantive task to identify, in general, which information flows are relevant to understanding those intentional processes. This is to acknowledge that the reference class problem cannot be resolved with just information-theoretic tools. Recognizing the extreme heterogeneity of Nomic regularities and attendant initial conditions, a Nomic and Factive notion of natural information can serve this broad explanatory goal. I define Nomic, Factive Information (NFI) as follows:

> **NFI:** b's being in state N carries the information that
> s is F if and only if (i) s is F, and (ii) the covariation of
> b's being in state N with s's being in state F is Nomic.

NFI is similar to DT, but it is distinct in two important ways. First, in incorporates the notion of Nomicity I articulated above, and therefore does not exhibit DT's strictness problem. Second, NFI makes no explicit reference to probability. The requirement of Factivity entailed by DT's condition (i) can be readily expressed without mentioning probability, as I have just done in NFI's conditions (i). I suspect that Dretske's appeal to conditional probabilities (of 1) has been a cause of confusion in this discourse, because it suggests that the most straightforward way to address DT's strictness problem is to lower the required probability to something less than 1. As I argued, DT's reliance on Laws of Nature is really what leads to the strictness problem.

Because NFI makes no explicit reference to probabilities, one might worry that this notion of information has nothing to do with probability or uncertainty. That would be an incongruous result, given the Shannon-Weiner-inspired roots of this information-based approach to explaining intentional phenomena. However, NFI is still essentially a matter of uncertainty in that the *measure* of information carried by some event depends on what happens in counterfactual scenarios. That is, given that N carries information that s is F, the *amount* of this information depends on how probable it is that s would have been F (rather than G, H, etc) in the case that b were not N (but rather O, P, etc.) If s's being F is quite common across all possibilities for b, then b's being N does not carry much information about s's being F. For instance, something indicating that a friend has *not* won the lottery is a signal with a meager sum of information, whereas a signal indicating they did win would have a much higher measure of information.

Having a measure of the strength of informational relations is important; one wants to be able to distinguish between signals that are informative as to the same fact but to different degrees. A signal that a friend did not win the lottery does not say how many tickets they bought, and one might want to treat the friend differently depending on how many tickets they bought. So probabilities are important to information, but this is quite different from treating information content itself as probabilistic. Nomic, Factive Information can be given such a formal measure - it coheres with the Shannon-Weiner theory just fine. On the view I have defended, probabilities decide how much information is transmitted between signals in a given channel, and the existence of alternative possibilities (counterfactuals) that are probabilistically related is essential to all information signals, but the content of an information signal is not probabilistic - the content is fixed by what the Nomically related facts are.

5   Conclusion

Natural information is an indispensable idea for much theoretical work in science; one can hardly do without the idea that some events reduce uncertainty as to the occurrence of other events just as a matter of the inherent physical properties of these events. Natural information plays an especially prominent role in work that tries to shed light on the way physical systems can exhibit intentional processes - how one can have perceptions and motivations, for instance. Dretske's (1981) account of natural information is a useful jumping-off point for thinking about the philosophical issues involved, as its insights have been highly influential while certain of its shortcomings have been widely appreciated. However, the prevailing critical response to that account misunderstands exactly how it falls short, blurring together the problem of strictness with the problem of settling references classes. Settling references classes is necessary to draw a link between successful and unsuccessful ways of using information, and rejecting Dretske's Factivity condition does not help with this. Rather, the move to non-Factivity obfuscates and postpones the reference class problem, relying on the false assumption that locally similar states that play a similar explanatory role in behavior must carry the same information.

I have argues that an adequate understanding of natural information requires two main modifications to Dretske's (1981) view. First, carefully considering what it is for a correlation between physical events to be non-accidental, one should drop DT's reliance on Laws of Nature in favor of the much more inclusive idea of Nomicity. Second, one should recognize that the reference class problem - which must be solved if one is to derive intentional content partly from informational content - demands resources from outside of information theory. This means recognizing a basic limit on what information-theoretic tools can do to tell us about the content of perceptions, motivations, and other intentional processes. Appropriately combined with some other approach to the reference class problem, Nomic Factive Information can support an understanding of naturally occurring intentionality.

References

[1] D. M. Armstrong. *What is a Law of Nature?* Cambridge University Press, 1983.

[2] Clint Ballinger. Initial conditions and the 'open systems' argument against laws of nature. *Metaphysica*, 9(1):17–31, 2008.

[3] Jon Barwise and Jerry Seligman. *Information Flow: The Logic of Distributed Systems*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1997.

[4] Helen Beebee and David Papineau. Probability as a guide to life. *Journal of Philosophy*, 94(5):217–243, 1997.

[5] Joseph Butler. *The Analogy of Religion, Natural and Revealed, to the Constitution and Course of Nature*. James, John and Paul Knapton, 1736.

[6] Nancy Cartwright. *The Dappled World: A Study of the Boundaries of Science*. Cambridge University Press, 1999.

[7] Sergio Chibbaro and Angelo Vulpiani. Compressibility, laws of nature, initial conditions and complexity. *Foundations of Physics*, 47(10):1368–1386, 2017.

[8] Fred Dretske. Laws of nature. *Philosophy of Science*, 44(2):248–268, 1977.

[9] Fred Dretske. *Knowledge the flow of information / Fred Dretske*. MIT Press Cambridge, Mass, 1st mit press ed. edition, 1981.

[10] Fred Dretske. Misrepresentation. In R. Bogdan, editor, *Belief: Form, Content, and Function*, pages 17–36. Oxford University Press, 1986.

[11] Chris Eliasmith. A new perspective on representational problems. *Journal of Cognitive Science*, 6:97–123, 2005.

[12] Mathias Frisch. Laws and initial conditions. *Philosophy of Science*, 71(5):696–706, 2004.

[13] Justin Garson. A generalized selected effects theory of function. *Philosophy of Science*, 84(3):523–543, 2017.

[14] Peter Godfrey-Smith. Indication and adaptation. *Synthese*, 92(2):283–312, 1992.

[15] H. Paul Grice. Meaning. *Philosophical Review*, 66(3):377–388, 1957.

[16] Paul E. Griffiths. Functional analysis and proper functions. *British Journal for the Philosophy of Science*, 44(3):409–422, 1993.

[17] Alan Hájek. The reference class problem is your problem too. *Synthese*, 156(3):563–585, 2007.

[18] Gilbert Harman. Problems with probabilistic semantics. In Alex Orenstein and Rafael Stern, editors, *Developments in Semantics*, pages 243–237. Haven, 1983.

[19] Pierre Jacob. Intentionality. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2019 edition, 2019.

[20] Daniel M. Kraemer. Natural probabilistic information. *Synthese*, 192(9):2901–2919, 2015.

[21] P. N. Kugler and M. T. Turvey. *Information, Natural Law, and the Self-Assembly of Rhythmic Movement*. Lawrence Erlbaum Associates, Inc., 1987.

[22] Marc Lange. Natural laws and the problem of provisos. *Erkenntnis*, 38(2):233–248, 1993.

[23] David K. Lewis. *Convention: A Philosophical Study*. Wiley-Blackwell, 1969.

[24] Ruth G. Millikan. *Language, Thought, and Other Biological Categories*. MIT Press, 1984.

[25] Ruth Garrett Millikan. What has natural information to do with intentional representation? In D. Walsh, editor, *Royal Institute of Philosophy Supplement*, pages 105–125. Cambridge University Press, 2001.

[26] Ruth Garrett Millikan. *Varieties of Meaning: The 2002 Jean Nicod Lectures*. MIT Press, 2004.

[27] Ruth Garrett Millikan. *Beyond Concepts: Unicepts, Language, and Natural Information*. Oxford University Press, 2017.

[28] Karen Neander. Functions as selected effects: The conceptual analyst's defense. *Philosophy of Science*, 58(2):168–184, 1991.

[29] Karen Neander. *A Mark of the Mental: A Defence of Informational Teleosemantics*. Cambridge, USA: MIT Press, 2017.

[30] Andrea Scarantino. Information as a probabilistic difference maker. *Australasian Journal of Philosophy*, 93(3):419–443, 2015.

[31] Andrea Scarantino and Gualtiero Piccinini. Information without truth. *Metaphilosophy*, 41(3):313–330, 2010.

[32] Claude E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27(3):379–423, 1948.

[33] Nicholas Shea. Consumers need information: Supplementing teleosemantics with an input condition. *Philosophy and Phenomenological Research*, 75(2):404–435, 2007.

[34] Brian Skyrms. *Signals: Evolution, Learning, and Information*. Oxford University Press, 2010.

[35] John Maynard Smith. The concept of information in biology. *Philosophy of Science*, 67(2):177–194, 2000.

[36] Michael Strevens. Explanation, abstraction, and difference-making. *Philosophy and Phenomenological Research*, 99(3):726–731, 2019.

[37] Bas C. van Fraassen. *Laws and Symmetry*. Oxford University Press, 1989.

[38] James Woodward. Laws: An invariance- based account. In Ott and Patton, editors, *Laws of Nature*. Oxford University Press, 2018.

[39] James Woodward. Physical modality, laws, and counterfactuals. *Synthese*, 197(5):1907–1929, 2020.