Studies in History and Philosophy of Science 41 (2010) 237–246

Contents lists available at ScienceDirect

Studies in History and Philosophy of Science

journal homepage: www.elsevier.com/locate/shpsa

Computation vs. information processing: why their difference matters to cognitive science

Gualtiero Piccinini[a], Andrea Scarantino[b]

[a] Department of Philosophy, University of Missouri St. Louis, 599 Lucas Hall (MC 73), One University Boulevard, St. Louis, MO 63121-4499, USA
[b] Department of Philosophy, Neuroscience Institute, Georgia State University, PO Box 4089, Atlanta, GA 30302-4089, USA

ARTICLE INFO

Keywords:
Computation
Information processing
Computationalism
Computational theory of mind
Cognitivism

ABSTRACT

Since the cognitive revolution, it has become commonplace that cognition involves both *computation* and *information processing*. Is this one claim or two? Is computation the same as information processing? The two terms are often used interchangeably, but this usage masks important differences. In this paper, we distinguish information processing from computation and examine some of their mutual relations, shedding light on the role each can play in a theory of cognition. We recommend that theorists of cognition be explicit and careful in choosing notions of computation and information and connecting them together.

When citing this paper, please use the full journal title *Studies in History and Philosophy of Science*

1. Introduction

Since the cognitive revolution, it has become commonplace that cognition involves both *computation* and *information processing*. Is this one claim or two? Is computation the same as information processing? The two terms are often used interchangeably, but this usage masks important differences. In this paper, we will distinguish information processing from computation and examine some of their mutual relations, shedding light on the role each can play in a theory of cognition. We will conclude by recommending that theorists of cognition be explicit and careful in choosing notions of computation and information and connecting them together.

One possible reason to assimilate 'computation' and 'information processing' is to mark a generic distinction between two approaches to cognitive science. On one side stands mainstream cognitivism, based on the manipulation of representations. On the other side, there are non-cognitivist approaches such as behaviorism and Gibsonian psychology, which reject mental representations. The thought may be that, by rebranding cognitivism as either computationalism or information-processing psychology, one is simply hinting at the divide between representational and anti-representational approaches. As we hope to make clear, this assimilation of computation, information processing, and the manipulation of representations does more harm than good.

Each of the central notions can be legitimately interpreted in different ways. Depending on what we mean by 'computation', 'information processing', and 'manipulation of representations', that something computes may or may not entail that it processes information, that it processes information may or may not entail that it manipulates representations, and behaviorism or Gibsonian psychology may or may not stand in opposition to either computationalism or information processing psychology.

Of course, one might explicitly stipulate that 'computation' and 'information processing' designate the same thing. As Smith (2002) points out, computation is sometimes *construed* as information processing. This tells us nothing about, and is in tension with, the independently established meanings of the terms—the meanings that are historically most influential and theoretically most important.

So we will set aside any stipulation that identifies computation and information processing. A useful starting point for understanding their relations is to reconstruct why and how the two notions were introduced in the sciences of mind and brain. Attending to this history can help us see where our intuitions about computation and information processing come from. Our ultimate objective

---

E-mail addresses: gpiccini@artsci.wustl.edu (G. Piccinini), ascarantino@gsu.edu (A. Scarantino)

0039-3681/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.shpsa.2010.07.012

is to move away from an unhelpful tug of war between opposing but similarly brute intuitions about computation and information processing, articulate clear distinctions between these notions, and assess which of them matter for cognitive science.

2. Historical preliminaries

The notions of computation and information came into the sciences of mind and brain from different places through different—though partially overlapping—routes. And they came to play different roles. Lacking room for a detailed historical treatment, we will limit ourselves to the following brief remarks.

The notion of computation came into cognitive science from mathematics, where a computation, in its main original sense, is an algorithmic process: a process that generates correct results by following an effective procedure. (Roughly, an effective procedure is an explicit step-by-step procedure guaranteed to produce the correct result for each relevant input.) This notion was made formally precise by Alan Turing and other logicians in the 1930s (Gödel, 1934; Church, 1936; Post, 1936; Turing, 1936–1937). A few years later, Warren McCulloch and Walter Pitts (1943) used Turing's formal notion of computation to characterize the activities of the brain. McCulloch and Pitts's computational theory had a large influence on computer design, artificial intelligence, and, eventually, the sciences of mind and brain (Piccinini, 2004a). Via McCulloch and Pitts, Turing's notion of computation became a way to recruit the tools of the logician (such as proof-theoretic methods) and those of the computer scientist (algorithms, computer programs, certain neural networks) in order to characterize the functionally relevant properties of psychological and neural processes.

By contrast, the notion of information came into cognitive science from control engineering (Rosenblueth et al., 1943; Wiener, 1948) and communication engineering (Shannon, 1948). For engineering purposes, information is what is transmitted by messages carried either within a system for purposes of feedback or across a physical channel for purposes of reproduction at a destination. Informal notions of information had been invoked in neurobiology as early as Edgar Adrian (1928) (cf. Garson, 2003), but no systematic formal analysis of the sense in which signals carry information had been provided prior to the mid-century. Norbert Wiener and Claude Shannon shared the insight that the transmission of information is fundamentally related to the reduction of uncertainty or entropy, and that the latter can be formally quantified.

Whereas Wiener (1948) focused on the role of information in the control of both mechanical and biological systems, Shannon (1948) was interested in the narrower question of how to transmit information efficiently across communication channels through proper encoding of the messages. Shannon's work immediately influenced psychologists such as George Miller, who put Shannon's technical notion of information at the foundation of a new science of mind (Miller & Frick, 1949; Miller, 1951; cf. Garner, 1988; Crowther-Heyck, 1999). Shannon's information theory entered neuroscience around the same time (MacKay & McCulloch, 1952). Measures of information inspired by Shannon have been used ever since to quantify various aspects of neural signals (Rieke et al., 1997; Dayan & Abbott, 2001). Shannon's theory also inspired philosophical theories of information (Dretske, 1981) and the broader informational semantics movement, which seeks to use information to naturalize mental content.

Right around the time they entered psychology and neuroscience, the notions of computation and information merged into what seemed an appealing synthesis. Roughly, the mind/brain was seen as a computer, whose function is to receive information from the environment, process it, and use it to control the body and perform intelligent actions (Wiener, 1948; McCulloch, 1949; Jeffress, 1951; Ashby, 1952; Bowden, 1953; Shannon & McCarthy, 1956; Miller, 1956; von Neumann, 1958; Miller et al., 1960; Feigenbaum & Feldman, 1963). Since then, computation and information processing have become almost inseparable—and often indistinguishable—in much literature on the mind and brain.

As it often happens, the marriage between computation and information processing is stormier than it may seem at first. The main source of trouble is the polysemy of the wedded concepts. They both have multiple, yet related meanings, which celebrants of the marriage ignore at their own peril. Our first task is to distinguish some importantly different notions of computation and information.

3. Computation

The notion of computation has been characterized in many ways. For present purposes, different notions of computation vary along two important dimensions. The first is how encompassing the notion is, that is, how many processes it includes as computational. The second dimension has to do with whether being the vehicle of a computation requires possessing semantic properties. Let's look at the first dimension first.

3.1. Digital computation

We use 'digital computation' for whatever notion is implicitly defined by the classical mathematical theory of computation, most famously associated with Turing. By this, we do not mean to appeal to a specific formalism, such as Turing machines. We mean to appeal to the class of formalisms of which Turing machines are an example and the kind of mathematical theorems that may be proven about them—including, of course, theorems about functions that are not computable by Turing machines (for an introduction, see Davis et al., 1994). We also do not mean to restrict ourselves to processes that follow algorithms. We include any process whose input–output relations can be characterized by the kind of mathematics originally developed by Turing and other computability theorists. These include the processes performed by many (though not all) kinds of neural networks.

There have been different attempts to explicate how the notion of digital computation applies to concrete physical systems. We will rely on the account one of us has developed in recent years (Piccinini, 2007a). Roughly, digital computation is the processing of strings of digits according to general rules defined over the digits. A digit in this sense need not represent a number—it is simply a component's state, whose type can be reliably and unambiguously distinguished from other types by the computing system.

The present notion of digital computation should not be confused with the notion of *classical* computation sometimes employed in cognitive science debates. To a first approximation, classical computation is digital computation performed in accordance with a fixed, step-by-step algorithm, perhaps in response to a representation of the algorithm internal to the system. Digital computation in the present sense is a broader notion, because it requires neither that the rules that define the computation be represented in the computing system, nor that the rules constitute a fixed algorithm followed step-by-step by the computing system. All that the rules need to do is specify what relationship obtains between the strings of digits constituting the inputs and the strings of digits constituting the outputs.

Many important distinctions may be drawn within digital computation, such as hard-wired vs. programmable, special purpose vs.

general purpose, and serial vs. parallel (cf. Piccinini, 2008a,b). Nevertheless, digital computation is the most restrictive notion of computation that we will consider here. It includes processes that follow ordinary algorithms and effective procedures, such as the computations performed by standard digital computers, as well as many types of connectionist processes, including those performed by standard backpropagation networks. Since digital computation is the notion that inspired the computational theory of cognition—in both its classical and connectionist variants—it is the most relevant notion for present purposes.[1]

3.2. Generic computation

Digital computation is traditionally contrasted with analog computation (for a detailed account of this contrast, see Piccinini, 2008b, Sect. 3.5). Analog computation, in turn, is a vague and slippery concept. The clearest notion of analog computation is that of computation by an analog computer (as defined in Pour-El, 1974). Roughly, an analog computer is a device whose function is to process continuous variables so as to solve certain systems of differential equations. It is worth remembering that shortly after McCulloch and Pitts argued that the brain performs digital computations, others offered the alternative theory that the brain performs analog computations, meaning, roughly, processes like those performed by analog computers (Gerard, 1951; see also Rubel, 1985).

The theory that the brain is literally an analog computer was never very popular, perhaps for the good reason that the main vehicles of neural processes appear to be trains of neuronal spikes, which do not look much like continuous variables (Dayan & Abbott, 2001).[2] In fact, they are made of all-or-none spikes, which are discontinuous events; this was the main reason why McCulloch and Pitts concluded, rightly or wrongly, that neural processes are digital computations.

In recent decades, many neuroscientists have started using the term ‘computation’ for the processing of neuronal spike trains (i.e. sequences of spikes produced by neurons in real time). The processing of neuronal spike trains by neural systems is often called ‘neural computation’. Whether neural computation is a form of digital computation, analog computation, or a type of process distinct from both digital and analog computation is a difficult question, which we cannot settle here. Instead, we will subsume digital computation, analog computation, and neural computation under the banner of ‘generic computation’.

We use ‘generic computation’ to designate *any process whose function is to manipulate medium-independent vehicles according to a rule defined over the vehicles*, where a medium-independent vehicle is such that all that matters for its processing are the differences between the values of different portions of the vehicle along a relevant dimension (as opposed to more specific physical properties, such as the vehicle’s material composition). Since a generic computation is medium-independent, it can be instantiated by many different physical mechanisms, so long as they possess a medium with an appropriate dimension of variation.[3]

For instance, digits are a kind of medium-independent vehicles: they can be implemented by many kinds of physical media (mechanical, electro-mechanical, electronic, etc.), but all that matters for the computations defined over them is that the relevant portions of such media belong to types that are unambiguously distinguishable by the processing mechanism. In other words, digital computers that manipulate different media may well perform the same (digital) computation.

By contrast, analog computers cannot always distinguish between any two portions of the continuous variables they manipulate—there is always a margin of measurement error. Since the system cannot distinguish unambiguously between two portions of a continuous variable, continuous variables and portions thereof are not digits or strings of digits. Nevertheless, all that matters for analog computations are the differences between the different portions of the variables being manipulated, to the degree that they can be distinguished by the system. Any further physical properties of the media that implement the variables are irrelevant to the computation. Like digital computations, analog computations operate on medium-independent vehicles.

Finally, current evidence suggests that the vehicles of neural processes are neuronal spikes and that the functionally relevant aspects of neural processes are medium-independent aspects of the spikes—primarily, spike rates (as opposed to any more concrete properties of the spikes). Thus, spike trains appear to be another case of medium-independent vehicle, in which case they qualify as proper vehicles for generic computations.

In conclusion, generic computation includes digital computation, analog computation, neural computation (which may or may not reduce to digital or analog computation), and perhaps more. Assuming that brains process spike trains and that spikes are medium-independent vehicles, it follows by definition that brains perform generic computations.

3.3. Semantic vs. non-semantic computation

So far, we have taxonomized different notions of computation according to how broad they are, namely, how many processes they include as computational. Now we will consider a second dimension along which notions of computation differ. Consider digital computation. The digits are often taken to be representations, because it is assumed that computation requires representation (Fodor, 1981).

One of us has argued at length that computation *per se*, in the sense implicitly defined by the practices of computability theory and computer science, does not require representation, and that any semantic notion of computation presupposes a non-semantic notion of computation (Piccinini, 2004b, 2008c). Meaningful words such as ‘avocado’ are both strings of digits and representations, and computations may be defined over them; nonsense sequences such as ‘2#r %h@’, which represent nothing, are strings of digits too, and computations may be defined over them just as well.

Although computation does not *require* representation, it certainly *allows* it. In fact, generally, computations *are* carried out over representations. For instance, almost without exception, the states manipulated by ordinary computers are representations.

To maintain generality, we will consider both semantic and non-semantic notions of computation. In a nutshell, semantic no-

---

[1] In contrasting the notion of classical computation with that of connectionist computation, we are following common usage (Fodor & Pylyshyn, 1988). But that contrast is a false one. Not only are many paradigmatic types of connectionist computation digital in the present sense, but some of them—such as those instantiated by McCulloch–Pitts nets (McCulloch & Pitts, 1943)—are classical too. Speaking more carefully, we may use the notion of digital computation to taxonomize theories of cognition into classical theories (which may or may not be connectionist), non-classical computational (connectionist) theories, and non-digital computational (connectionist) theories (Piccinini, 2008a).Incidentally, notice that Fodor and Pylyshyn’s notion of classical system is even more restrictive than the one we employ here, since for them, classical computations are defined over interpreted symbols and are sensitive to the combinatorial syntax of the symbols.

[2] The claim that the brain is an analog computer is ambiguous between two interpretations. On the literal interpretation, which we are considering in the text, ‘analog computer’ is given Pour-El’s precise meaning (Pour-El, 1974). On a looser interpretation, ‘analog computer’ refers to some not well specified class of (presumably non-digital) systems. This looser interpretation of ‘analog computer’ is not uncommon, but we find it misleading because prone to being confused with the literal interpretation.

[3] For this notion of medium-independence, we are indebted to Garson (2003) and a comment by Oron Shagrir.

tions of computation define computations as operating over representations. By contrast, non-semantic notions of computation define computations without requiring that the vehicles being manipulated be representations.

Let us take stock. We have introduced the following notions of computation: digital computation (computation according to computability theory) and generic computation (which includes digital computation as a special case). Within each category, we may distinguish a semantic notion of computation, which presupposes that the computational vehicles are representations, and a non-semantic one, which doesn't.[4]

4. Information processing

In principle, instances of information processing can differ along two dimensions: the type of processing involved and the type of information involved. We wish to be tolerant on what counts as processing, putting no principled restriction of what types of processing may be relevant. To differentiate between different notions of information processing, the key term for our purposes is 'information'.

In the broadest sense, the production of information is contingent upon the reduction of uncertainty. Uncertainty, in turn, can be understood either as a mind-dependent property (e.g. Rob's uncertainty about election results) or as a mind-independent property (e.g. the uncertainty of a radioactive atom decaying tomorrow). Different notions of uncertainty serve different purposes.

Whatever notion of uncertainty we choose, we can think of signals—for example, footprints in the sand, alarm calls, words, and so on—as *informative* to the extent that they *reduce uncertainty*. The reduction of uncertainty may have effects (e.g. on beliefs, behaviors, etc.), the prediction and explanation of which is generally what motivates information-talk in the first place.

In what follows, we will distinguish between *semantic* and *non-semantic* notions of information. We begin by introducing Shannon's influential theory of non-semantic information, and then proceed to distinguish two varieties of semantic information.

4.1. Shannon information

We use 'Shannon information' to designate the notion of non-semantic information formalized by Shannon (1948).[5] Shannon developed his technical notion of information with a specific objective in mind. He was interested in an optimal solution to what he called the 'fundamental problem of communication' (ibid., p. 379), that is, the reproduction of messages from an information source to a destination. In Shannon's sense, any device that produces a message in a probabilistic manner can count as an information source.

Shannon's notion of message is not the usual one. In the standard sense, a message has semantic content, which is to say that there is something it stands for. Additionally, messages stand for their semantic contents arbitrarily, as opposed to, say, by virtue of a physical relation. For instance, there is no non-arbitrary reason why inscriptions with the shape 'apple' stand for apples.

Shannon's messages need not have semantic content at all, which is to say that they need not stand for anything. They don't even have to be strings of digits of finitely many types; on the contrary, they may be continuous variables. We will continue to discuss Shannon's theory using the term 'message' for historical accuracy, but the reader should keep in mind that the usual commitments associated with the notion of message do not apply.

The identity of a communication-theoretic message is fully described by two features: (a) it is a physical structure distinguishable from a set of alternative physical structures, and (b) it belongs to an exhaustive set of mutually exclusive physical structures selectable with well defined probabilities. As Shannon emphasized, the 'semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is *selected from a set* of possible messages' (ibid., p. 379).

Under these premises, to *communicate* a message produced at a source amounts to generating a second message at a destination that replicates the original message so as to satisfy a number of desiderata (accuracy, speed, cost, etc.). The sense in which the semantic aspects are irrelevant is precisely that messages carry information merely *qua* selectable, physically distinguishable structures associated with given probabilities. A nonsense message such as '2#r %h@' could in principle generate *more* Shannon information than a meaningful message such as 'avocado'.

To get an intuitive grip on this idea, consider an experiment described by a random variable taking values a_1 and a_2 with probabilities $p(a_1) = 0.9999$ and $p(a_2) = 0.0001$ respectively. Before the experiment takes place, outcome a_1 is almost certain to occur, and outcome a_2 almost certain not to occur. The occurrence of both outcomes generates information, in the sense that it resolves the uncertainty characterizing the situation before the experiment takes place (in the absence of uncertainty, e.g. when $p(a_1) = 1$ and $p(a_2) = 0$, no Shannon information can be produced).

Shannon's intuition was that the occurrence of a_2 generates *more information* than the occurrence of a_1, because it is *less expectable*, or *more surprising*, in light of the probability distribution. Therefore, if in light of this distribution '2#r %h@' is more surprising than 'avocado', it will carry more Shannon information, despite its meaninglessness.

Soon after its appearance, Shannon's theory started being used in many fields for which it had not been designed, including neuroscience and psychology, economics, cryptography, and physics. For instance, Shannon information is commonly used by neuroscientists to measure the quantity of information carried by neural signals about a stimulus and estimate efficiency of coding (what forms of neural responses are optimal for carrying information about stimuli) (Dayan & Abbott, 2001, Ch. 4).

What worried Shannon was mostly the association of semantic aspects to the programmatically *non-semantic* notion of information he had formulated. 'Workers in other fields', Shannon emphasized, 'should realize that the basic results of the subject are aimed in a very specific direction ... the hard core of information theory is, essentially, a branch of mathematics, a strictly deductive system' (Shannon, 1993, p. 462). In what follows, we will keep this lesson in mind, and carefully distinguish between Shannon information and semantic notions of information.

4.2. Semantic information

Suppose that a certain process leads to the selection of signals. Think, for instance, of how you produce letters of the alphabet when speaking. Non-semantic information in the sense of Shannon has to do with the uncertainty that characterizes the *selection process as a whole*. The non-semantic information generated by the selection of a particular letter is a function of the degree to which the signal-producing process is 'free' to select among alternatives. If, say, forty equiprobable letters are all selectable, a great deal of

---

[4] One terminological caveat. Later on, we will also speak of *semantic* notions of information. In the case of non-natural semantic information, by 'semantic' we will mean the same as what we mean here, that is, representational. A representation is something that can mis-represent, that is, may be unsatisfied or false. In the case of natural information, by 'semantic' we will mean something broader, which is not representational because it cannot mis-represent (see below).

[5] For a history of Shannon's theory of communication, see e.g. Pierce (1980).

selective freedom is present. If only two equiprobable letters are selectable, selective freedom is much reduced. In either case, the selection of any given letter is informative, in the non-semantic sense that it precludes the selection of its possible alternatives (thus, the selection of one out of forty equiprobable letters is more informative than the selection of one out of two equiprobable letters).

Semantic notions of information, on the contrary, have to do with what *a particular signal* stands for or means. To address the semantics of the signal, it is no longer sufficient to know which other signals might have been selected instead and with what probabilities. If a certain number of equiprobable letters are selectable, the selection of each one of them will generate semantic information depending on what that particular letter stands for. Whereas different equiprobable signals are all alike in their degree of *non-semantic informativeness*, they differ in their degree of *semantic informativeness*, insofar as the selection of each of them stands for something different. We call 'semantic information' the information a signal carries by reducing uncertainty *about* some state of affairs. In this case, semantic aspects are crucial: what information the signal carries is constitutively related to what the signal stands for.

The challenge for a theory of semantic information is to specify what grounds the 'standing for' relation. We distinguish between two main ways of spelling out the notion of 'standing for'. One will lead us to theories of *natural* semantic information. Roughly speaking, signals carry *natural* information by standing in an appropriate physical relation to what they are about. This is the sense in which smoke produced by fire carries information about fire. The other will lead us to theories of *non-natural* semantic information. Signals carry *non-natural* information by being arbitrarily associated with what they are about. This is the sense in which the word 'smoke' carries information about smoke.

4.2.1. Natural (semantic) information

When smoke carries natural information about fire, what is the basis for this informational link? In the contemporary literature, we find three main accounts of the physical connection that distinguishes informationally related from informationally unrelated events: *probabilistic* (Dretske, 1981; Millikan, 2004), *nomic* (Dretske, 1981; Fodor, 1990), and *counterfactual* (Loewer, 1983; Cohen & Meskin, 2006). These contributions find their primary source of inspiration in Dretske's (1981) attempt to develop a theory of semantic information on the basis of Shannon's (1948) communication theory.

What got Dretske interested in Shannon's (1948) theory was primarily that Shannon had treated information as an objective, physical commodity. Dretske wanted information to be objective in order to use it for the naturalization of intentional content. (Stampe, 1975, 1977, are earlier attempts in this direction). But Dretske also wanted information to capture the semantic aspects of communication that Shannon had explicitly excluded.

Accounting for these semantic aspects, Dretske suggested, requires constructing a theory of information around the idea that 'information is that commodity capable of yielding knowledge' (Dretske, 1981, p. 44). An important consequence follows from this approach: since knowledge entails truth, so does natural information. The intuition here is that since one cannot come to know that p from a signal unless p, the falsity of p makes it impossible for a signal to carry the natural semantic information that p.

Inspired by Shannon's (1948) theory, Dretske developed a number of constraints for the transmission of semantic natural information, and defended the following probabilistic definition as uniquely capable of satisfying them: '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)' (Dretske, 1981, p. 65).

This account of information was soon criticized for including reference to k, which designates the knowledge state of the receiver about what possibilities exist at the source.[6] Dretske's critics thought that defining information in terms of knowledge, a state already imbued with intentional content, prevented the use of information for the naturalization of intentional content. If information is to be the 'physical yeast and flour' we need to 'bake the mental cake', as Dretske had promised, we had better not find pieces of the cake already mixed in with the yeast and flour.

Theorists of natural information proposed two main strategies to circumvent this problem. The first is to identify the physical connection between informative signals and what they are about with a nomic correlation (Dretske, 1981; Fodor, 1990). Under this view, what is needed for smoke to carry information about fire is the presence of a nomic correlation between smoke and fire.

The second strategy is to identify the physical connection between informative signals and the states of affairs they are about with a counterfactual connection (Cohen & Meskin, 2006). Under this view, what is needed for smoke to carry information about fire is the (non-vacuous) truth of the following counterfactual: if fire had not occurred, smoke would not have occurred.

In the rest of our paper, we focus on nomic theories of natural information, which are currently the received view in informational semantics (for a discussion of the shortcomings of counterfactual theories of information, see Scarantino, 2008).

4.2.2. Non-natural (semantic) information

In ordinary language, we commonly talk about false information or misinformation. If someone were to whisper in your ear, 'Hey, your opponent has an ace', when in fact your opponent has a queen, you might say that you have received false information. There are two principal strategies to make sense of this way of talking.

On one hand, we could say that you have received what *purported to be* the information that your opponent has an ace, but deny that any actual information was received. Under this view, information entails truth: if your opponent does not have an ace, you cannot possibly receive the information that he does. This is the road taken by most philosophers of semantic information, including Dretske (1981), Millikan (2004), and Floridi (2005).

On the other hand, we could say that you have actually received the information that your opponent has an ace even though, as a matter of fact, he does not. This interpretation presupposes a kind of semantic information that, unlike natural information, does not entail truth.

We choose this second option, because we are interested in a taxonomy of actual uses of the term 'information', and the idea that one can receive information that happens to be false is an important part of ordinary information talk. More importantly, notions of information that do not entail truth are often presupposed by psychologists and computer scientists when they speak of information processing. We call the kind of semantic information that can be false 'non-natural information'.

The natural vs. non-natural distinction imports into the domain of information a distinction Grice (1957) drew between two notions of meaning.[7] The first kind of meaning is exemplified by a sentence such as 'those spots mean measles', which is true, Grice claimed, just in case the patient has measles. The second kind of

---

[6] Another prominent criticism was that Dretske's (1981) theory sets too high a bar for the transmission of information, because signals in natural environments rarely raise the probability of what they are about to 1.

[7] Another important historical antecedent is Peirce's (1931) work on types of sign.

meaning is exemplified by a sentence such as 'those three rings on the bell (of the bus) mean that the bus is full' (ibid., p. 85), which is true even if the bus in not full. Similarly, we will say that 'x carries natural information that p' entails that p, whereas 'x carries non-natural information that p' does not entail that p.

Bearers of natural information stand for what they are about by virtue of being physically connected to it. In the absence of the connection, no natural information is carried. For instance, spots of a certain kind carry natural information about measles because of the physical connection between having measles and having that kind of spot.

Bearers of non-natural information, by contrast, need not be physically connected to what they are about. A convention, as in the case of the three rings on the bell of the bus, suffices to establish a non-natural informational link between events. The convention may either be explicitly stipulated, as in the rings case, or emerge spontaneously, as in the case of the non-natural information attached to words in natural languages.

The notion of non-natural information is often used interchangeably with the notion of representation. One may say that those three rings on the bell *represent* that the bus is full. If the bus is not full, the three rings will provide a *false* representation (a misrepresentation), or, equivalently, they will carry *false* non-natural information. By contrast, if the bus is not full, nothing will be able to carry the natural information that it is, because natural information cannot be false.

What's the relation between natural and non-natural information? As mentioned above, one motivation for theorizing about natural information is to use it to naturalize intentional content. We have now introduced a notion of information that, by corresponding to Grice's non-natural meaning or representation, is already imbued with intentional content. By calling such information non-natural, we are not taking a stance on whether it can be naturalized. We are simply using Grice's terminology to distinguish between two important notions of semantic information. We remain neutral on the prospects of reducing one to the other.

In conclusion, the differences between the three notions of information we are focusing on can be exemplified as follows. Consider an utterance of the sentence 'I have a toothache'. It carries Shannon information just in case the production of the sequence *I-blank-h-a-v-e-blank-a-blank-t-o-o-t-h-a-c-h-e* can be modeled by a stochastic process that generates letters with certain probabilities. The same utterance carries natural information about having a toothache just in case the production of the sequence *I-blank-h-a-v-e-blank-a-blank-t-o-o-t-h-a-c-h-e* is nomically correlated with having a toothache. Carrying natural information about having a toothache entails having a toothache. Finally, the same utterance carries non-natural information just in case the sequence *I-blank-h-a-v-e-blank-a-blank-t-o-o-t-h-a-c-h-e* has a non-natural meaning in a natural or artificial language. Carrying non-natural information about having a toothache does not entail having a toothache.

5. Computation may or may not be information processing

Are the vehicles over which computations are performed *necessarily* bearers of information? Is information processing *necessarily* carried out by means of computation? Answering these questions requires matching computation and information in several of the senses we have distinguished.

We will argue that the vehicles over which computations are performed may or may not bear information and that information processing may or may not be carried out by means of computation. Yet, as a matter of contingent fact, usually the vehicles over which computations are performed do bear information.

5.1. Computation vs. processing Shannon information

Few cognitive scientists should care about whether computation is the processing of Shannon information.[8] This is because cognitive scientists are generally interested in computational vehicles that bear *semantic* information. Nevertheless, for completeness' sake, let's briefly look at why computation need not be the processing of Shannon information, and why the processing of Shannon information need not be computation.

The notion of digital computation does not require that the selection of the computation vehicles produce Shannon information. Here is why. The vehicles of digital computation are strings of digits. The selection of digits produces Shannon information insofar as the selection process is characterized by uncertainty. The Shannon information generated by the selection of each specific digit is inversely related to its probability: the less expected the selection of a given digit is, the more Shannon information its selection produces. The selection of digits can also generate Shannon information in the context of a communication channel, which consists of two ensembles of statistically correlated events. If the selection of digits changes the degree to which some other events are expectable, it will produce Shannon information relative to them. By these criteria, the selection of digits need not produce Shannon information.

First, digits need not be statistically correlated with anything else in order to undergo a digital computation. Consider a computer performing multiplication. Before it starts computing, it has input data and an initial internal state, both of which may be characterized as strings of digits. Input data and initial internal states may correlate in some interesting way with some other variable, but they don't have to. In the typical case, they are simply inserted in the computer by its user. More importantly, what matters for a computation to be well defined and successfully carried out is that the digits be there, not that they correlate with anything else. Thus, the selection of a computation's digits need not produce Shannon information about any correlated events.

Second, a digit may well have probability 1 of being selected. Nothing in the notion of digital computation prevents that a computation be performed on a digit that was certain to be selected by some selection mechanism. In such a case, the selection of that digit does not produce Shannon information, because there is no uncertainty to be reduced concerning which digit is selected. Furthermore, computation may be deterministic or indeterministic, but most computational steps are deterministic. Deterministic computations generate new digits with probability 1.[9] Therefore, most computationally produced digits generate no Shannon information. Thus, the selection of a computation's digits need not generate Shannon information.

In summary, since computation is a process defined over digits and the selection of digits need not produce any Shannon information, digital computing does not entail the processing of Shannon information.

(To be sure, computing generally requires the transmission of Shannon information across communication channels between the system's components. Shannon's information theory can be

---

[8] This being said, cognitive scientists do care about the efficient encoding of neural signals, and may use Shannon's theory to analyze some aspects of the reproduction of signals across communication channels.

[9] We are following Shannon in taking the probabilities in question to be objective. Of course, if one were to replace objective probabilities with subjective ones, even a deterministic computation may well generate Shannon information.

usefully applied to studying the efficiency of computer codes (e.g. for the purpose of data compression, eliminating redundancies in a code, or devising the most efficient code for maximizing the average rate of information communicated) as well as the transmission of signals within computers (noise minimization).)

Conversely, processing Shannon information may or may not be done by means of digital computation. First, for it to be done by digital computing, Shannon information must be produced and carried by strings of digits. But Shannon information can be produced and carried by continuous signals, which may be processed by analog means. Second, even when the bearers of Shannon information are digits, there exist forms of processing of such digits other than digital computation. Just to give an example, a digital-to-analog converter transforms digits into analog signals.

In conclusion, digital computation may or may not entail the processing of Shannon information, and Shannon information may or may not be processed by means of digital computation.

The case we have made for digital computation can be extended to generic computation, which includes, besides digital computation, analog computation and neural computation. Nothing in the notions of analog computation or neural computation mandates that the selection of the pertinent signals produce Shannon information. Thus, generic computation may or may not entail the processing of Shannon information.

The main difference from the previous case is that now, Shannon information processing is necessarily a form of generic computation. As we defined it, generic computation is the functional manipulation of any medium-independent vehicle. Information is a medium-independent notion, in the sense that whether a vehicle generates Shannon information does not depend on its specific physical properties, but rather on its probability of occurrence. Thus, generic computation is broad enough to encompass the processing of Shannon information. In other words, if a vehicle carries Shannon information, its processing is a generic computation.[10]

5.2. Computation vs. processing natural information

Many neuroscientists use ‘computation’ and ‘information processing’ interchangeably. What they generally mean by ‘information processing’, we submit, is the processing of natural (semantic) information carried by neural responses. It is thus important to examine whether and in what sense computation is the processing of natural information.

5.2.1. Digital computation

The notion of digital computation does not require that the computation vehicles bear natural information—or more precisely, that they carry natural information *about the* computing system’s *environment*. At least for deterministic computation, there is always a nomic correlation between later states and earlier states of the system. Furthermore, any computational input and initial state will have causes with which it nomically correlates. In this sense, (deterministic) digital computation entails natural information processing. But the natural information being processed need not be about the external environment of the organism, which is what theorists of cognition are most interested in. In this section, we will focus on whether digital computation entails the processing of natural information about the system’s environment.

Granted, natural information is virtually ubiquitous—it is easy enough to find nomically underwritten correlations between physical variables. This being the case, the digits present within a digital computer may well carry natural information about cognitively interesting sources, such as environmental stimuli that the computer responds to. Consider a car’s computer, which receives and responds to natural information about the state of the car. The computer uses feedback from the car to regulate fuel injection, ignition timing, speed, and so on. In such a case, a digital computation will be a case of natural information processing.

But digits—the vehicles of digital computation—need not be nomically correlated with anything in a system’s environment in order for a digital computation to be performed over them: a digital computation may or may not constitute the processing of natural information. If the computation vehicles *do* carry natural information, this may be quite important. In our example, that certain digits carry natural information about the state of the car explains why the car’s computer successfully regulates fuel injection, ignition timing, and so on. Our point is simply that digital computation does not entail the processing of natural information.

This point is largely independent of the distinction between semantic and non-semantic digital computation. A semantic digital computation is generally defined over the *non*-natural information carried by the digits—independently of whether they carry natural information or what specific natural information they may carry. Furthermore, it is quite possible for a digital computation to yield an incorrect output, which misrepresents the outcome of the operation performed. As we have seen, natural information is necessarily true. Thus, even semantic digital computation does not entail the processing of natural information.

At this stage, someone might reply that non-natural information reduces to natural information—plus, perhaps, some other naturalistic ingredients. Several efforts have been made to this effect (Dretske, 1981, 1988; Fodor, 1990; Barwise & Seligman, 1997; Millikan, 2004). According to these theories, non-natural information is, at least in part, natural information. If this is right, then at least semantic digital computation entails the processing of natural information.

But we take seriously the possibility that the antecedent is false. Theories according to which non-natural information reduces to natural information are one family of theories among others. There are other theories according to which either non-natural information is irreducible, or it reduces to things other than natural information (e.g. Harman, 1987). Since the project of reducing non-natural information to natural information may or may not go through, semantic digital computation may or may not entail the processing of natural information. At this stage of the game, it cannot be simply assumed that semantic digital computation is the processing of natural information.

What about the converse claim: that the processing of natural information must be carried out by means of digital computation? Again, it may or may not be. It depends on whether the natural information is digitally encoded and how it is processed. Natural information may be encoded by continuous variables, which are not the vehicles of digital computation. Or it may be encoded by digits, but the processing of such digits may consist in something other than digital computation, as in the conversion of digital into analog signals. In both of these cases, non-natural information is processed, but not by means of digital computation.

In sum, digital computation vehicles need not carry natural information, and natural information processing need not be performed by digital computation.

5.2.2. Generic computation

Generic computation may or may not be the processing of natural information. This case is analogous to the previous one. We already established it in the case of digital computation, which is a special case of generic computation. It is just as easy to establish

---

[10] In this section, we have not discussed the distinction between semantic and non-semantic notions of computation, as it makes no difference to our present concerns.

it for other kinds of generic computation, such as analog computation and neural computation (i.e. the processing of neuronal spike trains). Nothing in the notions of analog computation or neural computation mandates that the vehicles being manipulated carry natural information. More generally, nothing in the notion of manipulating medium-independent vehicles mandates that such vehicles carry natural information.

The main difference from the previous case is that the converse does not hold. Now, natural information processing is necessarily a form of generic computation, because, again, the notion of generic computation was intentionally left broad enough to encompass all the relevant processes.

5.3. Computation vs. processing non-natural information

Non-natural information is a central notion in our discourse about minds and computers. We attribute conventional semantic contents to each other's minds. We do the same with the digits manipulated by computers. We will now examine whether and in what sense computation is the processing of non-natural information.

5.3.1. Digital computation

Digital computation may or may not be the processing of non-natural information, and vice versa. This case is analogous to that of natural information, with one major difference. In the case of *semantic* digital computation, any reasonable notion of semantic content turns digital computation into non-natural information processing. (Again, we are neutral on whether such 'non-natural' information may be naturalized.) As a matter of fact, virtually all digital computing conducted in artifacts *is* information processing, at least in the sense that it processes non-natural information.

But the main question we are trying to answer is whether *necessarily*, digital computation is the processing of non-natural information. It *is* if we define a notion of digital computation that entails that the digits are representations. But as pointed out in Section 3.3, nothing in the notion of digital computation used by computer scientists and computability theorists mandates such a definition. So, in the theoretically most important sense of digital computation—digital computation as implicitly defined by the practices of computability theory and computer science—digital computation need not be the processing of non-natural information.

Conversely, the processing of non-natural information need not be carried out by means of digital computation. It all depends on whether the non-natural information is digitally encoded and how it is processed. The information may be encoded digitally, or by continuous variables, or by other vehicles. And even when encoded digitally, it may be manipulated by digital computation or by other processes, such as in a digital-to-analog conversion.

In conclusion, digital computation vehicles need not carry non-natural information (though they do if the notion of digital computation is semantic), and the processing of non-natural information need not be a digital computation.

5.3.2. Generic computation

Generic computation may or may not constitute the processing of semantic information, for similar reasons. Nothing in the notions of analog computation or neural computation mandates that the vehicles being manipulated carry non-natural information. They could be entirely meaningless and the computations would proceed just the same. More generally, nothing in the notion of manipulating medium-independent vehicles mandates that such vehicles carry non-natural information. Again, the converse does not hold. Semantic information processing must be done by means of generic computation, for this notion of computation is broad enough to encompass the relevant processes.

6. Why the difference between computation and information processing matters to cognitive science

As our analysis demonstrates, 'information processing' and 'computation' are Protean terms. As soon as we clarify their multiple meanings—as we should prior to employing these terms in a rigorous theory of cognition—we realize that the mutual relations between computation and information processing are far from obvious. The alleged synonymy between computation and information processing, it turns out, rests either on unfruitful stipulation or on disregarding important distinctions between the two notions.

If this is right, why are the two notions used interchangeably so often, without a second thought? We suspect the historical reason for this conflation goes back to the cybernetic movement's effort to blend Shannon's information theory with Turing's computability theory (as well as control theory). Cyberneticians did not clearly distinguish either between Shannon information and semantic information or between semantic and non-semantic computation. Still, at least initially, they were fairly clear that information and computation played distinct roles within their theory. Their idea was that organisms and automata contain control mechanisms: information is transmitted within the system and between system and environment, and control is exerted by means of digital computation (or perhaps by means of analog computation, another special case of generic computation).

Then the waters got muddier. When the cybernetic movement became influential in psychology, AI, and neuroscience, 'computation' and 'information' became ubiquitous buzzwords. Many people accepted that computation and information belong together in a theory of cognition. After that, they stopped paying attention to the differences between the two, with the unfortunate consequence that confusions about computation started piling up with confusions about information. To set the record straight and make some progress, we must get clearer on the independent roles computation and information can fulfill in a theory of cognition.

The notion of digital computation was imported from computability theory into neuroscience and psychology primarily for two reasons: first, it seemed to provide the right mathematics for modeling neural activity; second, it inherited mathematical tools (algorithms, computer program, formal languages, logical formalisms, and their derivatives, including many types of neural networks) that appeared to capture some aspects of cognition. Whether these reasons are enough to establish that cognition is digital computation is a difficult question, which lies outside the scope of this essay.

The theory that cognition is computation became so popular that it progressively led to a stretching of the operative notion of computation. In many quarters, especially neuroscientific ones, the term 'computation' is used, more or less, for whatever internal processes explain cognition. We have included this notion under the rubric of 'generic computation'. Unlike 'digital computation', which stands for a mathematical apparatus in search of applications, 'neural computation' is a label in search of a theory. Of course, the theory is quite well developed by now, as witnessed by the explosion of work in theoretical and computation neuroscience over the last decades (O'Reilly & Munakata, 2000; Dayan & Abbott, 2001). The point is that such a theory need not rely on a previously existing and independently defined notion of computation, such as 'digital computation' or even 'analog computation' in its most straightforward sense.

By contrast, the various notions of information have entirely distinct roles to play. By and large, they serve to make sense of how organisms keep track of ecologically relevant events in their environments and produce behaviors accordingly. Shannon information can serve to address quantitative problems of efficiency of communication in the presence of noise, including communication between the environment and the nervous system. Natural information can serve to give specific semantic content to particular states or events. This may include cognitive or neural events, which are often nomically correlated with events occurring in the environment. Finally, non-natural information can serve to characterize the conventional meaning of concepts, words, and the thoughts and sentences they constitute.

Whether cognitive or neural events fulfill all or any of the job descriptions of computation and information processing is in part an empirical question and in part a conceptual one. It's a conceptual question insofar as we can mean different things by 'information' and 'computation', and insofar as there are conceptual relations between the various notions. It's an empirical question insofar as, once we fix the meanings of 'computation' and 'information', the extent to which computation and the processing of information are coinstantiated in the brain depends on the empirical facts of the matter.

In this essay, we have charted the main notions of computation and information that are relevant in cognitive science as well as their conceptual relations (or lack thereof). To wit, we have found that computation vehicles may or may not carry information; in turn, information processing may or may not be done by means of (digital) computing.

To be sure, cognition involves generic computation. This is a trivial conclusion, however, because we defined 'generic computation' to include any type of process that manipulates medium-independent vehicles. That cognition involves such vehicles is uncontroversial. It would be more interesting if cognition turned out to involve digital computation, which is defined independently of the theory of cognition and is the notion that initially inspired the computational theory of cognition. But the question of whether cognition involves digital computation, or some other nontrivial notion of computation, has yet to be settled.

Regardless of what kind of computation cognition may involve, it doesn't follow that cognition involves the processing of information. In fact, there are theories according to which cognition involves computation (or at least *may* involve computation), but it does not involve the processing of information (Stich, 1983; Ramsey, 2007). More precisely, what these authors reject is the role of representation in a theory of mind. They don't explicitly reject the notion of information. But notice that the notion of non-natural information is fully representational (it allows for representational error), and even the notion of natural information is often used by scientists and philosophers to underwrite a notion of representation (Dretske, 1981, 1988).[11] Theories according to which cognition involves computation but not representation have occasionally been accused of incoherence, on the grounds that computation (allegedly) requires representation. On the contrary, it is a corollary of our account that these theories are perfectly coherent. Whether they are correct, of course, is another story.

It is independently plausible that cognition involves information processing. After all, natural information is carried by neural and cognitive events by virtue of nomic correlations with environmental events, and such correlations are commonplace (O'Reilly & Munakata, 2000; Dayan & Abbott, 2001). Moreover, some aspects of such correlations may be measured by means of Shannon information. Finally, the concepts possessed and the words spoken by human cognizers are generally individuated in part by the non-natural information they carry. That human cognition involves the processing of non-natural information is not only commonsense; it is also presupposed by many mainstream psychological theories. Whether non-natural information reduces to natural information and other naturalistic ingredients is an important and difficult question, which we cannot take up here.

Ok, but do these distinctions really matter? Why should a cognitive theorist care about the differences between computation and information processing? The main theoretical advantage of keeping them separate is to appreciate the independent contributions they can make to a theory of cognition. Conversely, the main cost of conflating computation and information processing is that the resulting mongrel concept may be too vague and heterogeneous to do all the jobs that are required of it.

Whenever it is unclear which notions (if any) of information and computation different theorists employ and how they are related, cross-purpose talk lurks at every corner. It will be difficult to compare and contrast theories; non-equivalent theories may be taken to be equivalent and equivalent theories may be taken to be non-equivalent. The notion of non-natural information may be used without realizing that it requires naturalization in order to be part of a truly naturalistic theory of cognition. By the same token, the notion of natural information may be used without realizing that, unlike non-natural information, it entails truth. Making progress in understanding cognition requires clarity and precision on what kind of computation and information processing are involved in cognition, and on whether and how the two are related. If they are mutually related, it is important to understand whether they are related conceptually, because of how we use the notions of computation and information processing, or whether they are related in ways that constitute an empirical discovery about cognition.

The conflation of computation with information processing leads to fallacious arguments. For instance, we have become accustomed to arguments to the effect that cognition does not involve computation because cognition does not involve representations (e.g. van Gelder, 1995). Aside from the implausibility of the premise, the conclusion simply doesn't follow. Computationalism can survive the rejection of representationalism.

So one major consequence of our account is that the common view that cognition involves computation simply because it processes information is liable to generate serious mistakes. There are two important senses—corresponding to the two notions of natural and non-natural information—in which, plausibly, cognition involves information processing. Neither of these senses entails that cognition is computation in the historically and theoretically most significant sense of the term, that is, digital computation (which includes connectionist computation in the most straightforward sense). Thus, cognition may well involve the processing of information. Whether this information is processed by means of (digital) computation remains an open question.[12]

Acknowledgements

Many thanks to those who commented on this paper, including Neal Anderson, Nir Fresco, Mark Sprevak, the referees, our audience at the 2008 conference on Computation in Cognitive Science, and participants in the St. Louis Philosophy Reading Lunch. Piccinini was supported in part by a University of Missouri Research Board Award. A revised and expanded version of some portions of this article will appear in G. Piccinini and A. Scarantino, ''Information Processing, Computation, and Cognition,'' forthcoming in the *Journal of Biological Physics* (2010).

---

[11] Ramsey (2007) argues that this is not a genuine notion of representation. We are not persuaded, but we lack the space to do justice to this topic.

[12] One of us believes the answer is likely to be negative (Piccinini, 2007b).

References

Adrian, E. D. (1928). *The basis of sensation: The action of the sense organs*. New York: Norton.

Ashby, W. R. (1952). *Design for a brain*. London: Chapman and Hall.

Barwise, J., & Seligman, J. (1997). *Information flow: The logic of distributed systems*. Cambridge: Cambridge University Press.

Bowden, B. V. (1953). *Faster than thought*. London: Sir Isaac Pitman.

Church, A. (1936). An unsolvable problem in elementary number theory. *The American Journal of Mathematics, 58*, 345–363.

Cohen, J., & Meskin, A. (2006). An objective counterfactual theory of information. *Australasian Journal of Philosophy, 84*, 333–352.

Crowther-Heyck, H. (1999). George E. Miller, language, and the computer metaphor of mind. *History of Psychology, 2*, 37–64.

Davis, M. D., Sigal, R., & Weyuker, E. J. (1994). *Computability, complexity, and languages*. Boston: Academic.

Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience. Computational and mathematical modeling of neural systems*. Cambridge, MA: MIT Press.

Dretske, F. (1981). *Knowledge and the flow of information*. Cambridge, MA: MIT Press.

Dretske, F. (1988). *Explaining behavior*. Cambridge, MA: MIT Press.

Feigenbaum, E. A., & Feldman, J. (1963). *Computers and thought*. New York: McGraw-Hill.

Floridi, L. (2005). Is semantic information meaningful data? *Philosophy and Phenomenological Research, 70*, 351–370.

Fodor, J. A. (1981). The mind–body problem. *Scientific American, 244*, 114–123.

Fodor, J. A. (1990). *A theory of content and other essays*. Cambridge, MA: MIT Press.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*, 3–71.

Garner, W. R. (1988). The contribution of information theory to psychology. In W. Hirst (Ed.), *The making of cognitive science: Essays in honor of George A. Miller* (pp. 19–35). Cambridge: Cambridge University Press.

Garson, J. (2003). The introduction of information into neurobiology. *Philosophy of Science, 70*, 926–936.

Gerard, R. W. (1951). Some of the problems concerning digital notions in the central nervous system. In H. von Foerster, M. Mead, & H. L. Teuber (Eds.), *Cybernetics: Circular causal and feedback mechanisms in biological and social systems. Transactions of the Seventh Conference* (pp. 11–57). New York: Macy Foundation.

Gödel, K. (1934). On undecidable propositions of formal mathematical systems. In M. Davis (Ed.), *The undecidable* (pp. 41–71). Ewlett, NY: Raven.

Grice, P. (1957). Meaning. *The Philosophical Review, 66*, 377–388.

Harman, G. (1987). (Nonsolipsistic) conceptual role semantics. In E. LePore (Ed.), *New directions in semantics* (pp. 55–81). London: Academic Press.

Jeffress, L. A. (Ed.). (1951). *Cerebral mechanisms in behavior*. New York: Wiley.

Loewer, B. (1983). Information and belief. *Behavioral and Brain Sciences, 6*, 75–76.

McCulloch, W. S. (1949). The brain as a computing machine. *Electrical Engineering, 68*, 492–497.

McCulloch, W. S., & Pitts, W. H. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics, 7*, 115–133.

MacKay, D., & McCulloch, W. S. (1952). The limiting information capacity of a neuronal link. *Bulletin of Mathematical Biophysics, 14*, 127–135.

Miller, G. A. (1951). *Language and communication*. New York: McGraw-Hill.

Miller, G. A. (1956). The magical number seven, plus or minus a two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

Miller, G. A., & Frick, F. (1949). Statistical behavioristics and sequences of responses. *Psychological Review, 56*, 311–325.

Miller, G. A., Galanter, E. H., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt.

Millikan, R. G. (2004). *Varieties of meaning*. Cambridge, MA: MIT Press.

O'Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience. Understanding the mind by simulating the brain*. Cambridge, MA: MIT Press.

Peirce, C. S. (1931). *Collected papers*. Cambridge, MA: Harvard University Press.

Piccinini, G. (2004a). The first computational theory of mind and brain: A close look at McCulloch and Pitts's 'Logical calculus of ideas immanent in nervous activity'. *Synthese, 141*, 175–215.

Piccinini, G. (2004b). Functionalism, computationalism, and mental contents. *Canadian Journal of Philosophy, 34*, 375–410.

Piccinini, G. (2007a). Computing mechanisms. *Philosophy of Science, 74*, 501–526.

Piccinini, G. (2007b). Computational explanation and mechanistic explanation of mind. In M. De Caro, F. Ferretti, & M. Marraffa (Eds.), *Cartographies of the mind: Philosophy and psychology in intersection* (pp. 23–36). Dordrecht: Springer.

Piccinini, G. (2008a). Some neural networks compute, others don't. *Neural networks, 21*, 311–321.

Piccinini, G. (2008b). Computers. *Pacific Philosophical Quarterly, 89*, 32–73.

Piccinini, G. (2008c). Computation without representation. *Philosophical Studies, 137*, 205–241.

Pierce, J. R. (1980). *An introduction to information theory*. New York: Dover.

Pour-El, M. B. (1974). Abstract computability and its relation to the general purpose analog computer (some connections between logic, differential equations and analog computers). *Transactions of the American Mathematical Society, 199*, 1–28.

Post, E. (1936). Finite combinatorial processes: Formulation I. *Journal of Symbolic Logic, 1*, 103–105.

Ramsey, W. M. (2007). *Representation reconsidered*. Cambridge: Cambridge University Press.

Rieke, F., Warland, D., de Ruyter van Steveninck, R., & Bialek, W. (1997). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.

Rubel, L. A. (1985). The brain as an analog computer. *Journal of Theoretical Neurobiology, 4*, 73–81.

Scarantino, A. (2008). Shell games, information and counterfactuals. *Australasian Journal of Philosophy, 86*, 629–634.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*, 379–423, 623–656.

Shannon, C. E. (1993). The bandwagon. In idem, *Claude Elwood Shannon: Collected papers* (N. J. A. Sloane, & A. D. Wyner, Eds.) (p. 462). New York: IEEE Press.

Shannon, C. E., & McCarthy, J. (Eds.). (1956). *Automata studies*. Princeton, NJ: Princeton University Press.

Smith, B. C. (2002). The foundations of computing. In M. Scheutz (Ed.), *Computationalism: New directions* (pp. 23–58). Cambridge, MA: MIT Press.

Stampe, D. (1975). Show and tell. In B. Freed, A. Marras, & P. Maynard (Eds.), *Forms of representation: Proceedings of the 1972 Philosophy Colloquium of the University of Western Ontario* (pp. 221–245). Amsterdam: North-Holland.

Stampe, D. (1977). Toward a causal theory of linguistic representation. In P. French, T. Uehling, & H. Wettstein, (Eds.), *Studies in the philosophy of language* (pp. 42–63). Midwest Studies in Philosophy, 2. Minneapolis: University of Minnesota Press.

Stich, S. (1983). *From folk psychology to cognitive science*. Cambridge, MA: MIT Press.

Turing, A. M. (1936–1937). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, *Series 2*, *42*, 230–265; *43*, 544–546.

van Gelder, T. (1995). What might cognition be, if not computation? *The Journal of Philosophy, XCII*, 345–381.

von Neumann, J. (1958). *The computer and the brain*. New Haven, CT: Yale University Press.

Rosenblueth, A., Wiener, N., & Bigelow, J. (1943). Behavior, purpose and teleology. *Philosophy of Science, 10*, 18–24.

Wiener, N. (1948). *Cybernetics: Or control and communication in the animal and the machine*. Cambridge, MA: MIT Press.