# Evidence for Information Processing in the Brain

## Marc Burock

**Abstract**

Many cognitive and neuroscientists attempt to assign biological functions to brain structures. To achieve this end, scientists perform experiments that relate the physical properties of brain structures to organism-level abilities, behaviors, and environmental stimuli. Researchers make use of various measuring instruments and methodological techniques to obtain this kind of relational evidence, ranging from single-unit electrophysiology and optogenetics to whole brain functional MRI. Each experiment is intended to identify brain function. However, seemingly independent of experimental evidence, many cognitive scientists, neuroscientists, and philosophers of science assume that the brain processes information as a scientific fact. In this work we analyze categories of relational evidence and find that although physical features of specific brain areas selectively covary with external stimuli and abilities, and that the brain shows reliable causal organization, there is no direct evidence supporting the claim that information processing is a natural function of the brain. We conclude that the belief in brain information processing adds little to the science of cognitive science and functions primarily as a metaphor for efficient communication of neuroscientific data.

**Keywords**: brain function; cognitive science; experiments; information theory; neuroimaging; neuroscience evidence; philosophy of science

## 1. Introduction

Many of us believe that the brain processes information. Bechtel and Richardson (2010), as philosophers of cognitive science, consider it uncontroversial that cognitive scientists involved in neuroimaging research believe that "the brain contains some regions that are specialized for processing specific types of information" (p. 241). Neuroscientists too claim that "the principle function of the central nervous system is to represent and transform information" (deCharms & Zador 2000, p. 613). Given such wide-spread acceptance of a belief, it is appropriate to ask for the justification of this belief. If the justification is empirical and experimental, then we should look to the research reported by working scientists in the field; if it is metaphysical, then we should look to the arguments of philosophers and theoreticians.

We will no doubt discover both kinds of justification if we look for it. Yet we assume that cognitive scientists, when stating that the brain processes information, are primarily stating an empirical fact or a widely agreed-upon scientific proposition that is supported by a body of experimental evidence. Like the physicist who can back up the proposition 'protons have spin' with a presentation of the experimental evidence, we expect that the cognitive scientist should be able to do the same regarding a statement about the brain. If the cognitive scientist cannot do this, then the proposition is non-empirical or unscientific. We are not suggesting a definition of science or solving Popper's demarcation problem, but we are appealing to the belief that accepted scientific statements are associated with experimental evidence. Without associated evidence, a proposition cannot be scientific.

Our task here, however, is somewhat more involved than an objective review of the scientific literature. As we have learned from philosophers of science over the past century, "theory dominates the experimental work from its initial planning up to the finishing touches in the laboratory" (Popper 1959, p. 90). Hanson (1958) and Kuhn (1962) were among the first to direct our attention to the theory-ladenness of scientific observation. Brewster (2001) extended this position, arguing that the complete scientific process, which includes attention,

perception, data interpretation, memory, and scientific communication, is influenced by theory. Perhaps most relevant to our work is Popper's warning that "…observation statements and statements of experimental results, are always *interpretations* of the facts observed…they are *interpretations in the light of theories"* (italics in original, p. 90). Of course, none of this need imply scientific relativism, and relativism is not assumed in this work.

One may presume that only cognitive scientists are qualified to interpret the experimental evidence in the field. While an expert's assessment carries more weight than the non-involved observer, it is reasonable that anyone who takes time to understand the evidence and its methods of acquisition is in a position to construct an interpretation. The force of the interpretation should be based upon the reason of the argument and not only its source. Nonetheless, we have performed some of the types of experiments that we are now interpreting.

Cognitive scientific evidence, especially neuroimaging evidence, has been increasingly subjected to criticisms. To better demarcate our position, we highlight that we are not specifically arguing between distributed versus localized processing in the brain (Utel 2001; Hardcastle & Stewart 2002; Bunzl et al. 2010), and we are not pointing out the previously discussed technical-methodological limitations of brain assessing technologies (Logothetis 2008; Roskies 2007; Klein 2009). We do share with these authors the broader concern for interpretations of evidence in the field of cognitive science, and how theoretical assumptions influence interpretations of evidence, ultimately ending in statements made by cognitive scientists that carry the weight of scientific fact. These facts, in turn, are used by naturalistic philosophers of mind to constrain philosophical theory and argument.

## 2. Some cognitive science evidence types

The scientific statement that we will consider is Bechtel and Richardson's proposition "the brain contains some regions that are specialized for processing specific types of information," although, since we are not specifically arguing against localization of brain function, we will consider simultaneously the more general proposition P 'the brain processes information.' deCharms and Zador say that it is the *function* of the brain to process (represent and transform) information. There is no philosophical consensus on how to define a natural biological function, and we will assume that processing information is a natural function of the brain like pumping blood is a natural function of the heart.

One might expect our forthcoming analysis to be guided by a preliminary definition of information processing, but this will not be our method. Rather, we will directly consider the neurobiological evidence that scientists and philosophers presume to support information processing in the brain, and argue that this evidence does not yet justify propositions about the specific functioning of brain tissue, information processing or otherwise. We realize that this claim may initially (and perhaps finally) sound ridiculous to those practicing in the field of cognitive science. Of course, a manuscript on information processing would not be complete without reference to Claude Shannon's understanding of information, and we will show that his communication model, although immeasurably useful in modern technology and sometimes appropriately used in neuroscience, does not analogously apply to the communication between environment and brain.

Cognitive scientists, instead of claiming to discover the function of brain tissue, often speak of identifying cognitive *operations* with brain tissue or networks (Henson 2006; Bechtel 2008b). This difference does not substantially change our arguments. On Bechtel's view, cognitive operations are analogous to material operations, such as oxidizing a chemical substrate, where the cognitive substrates are mental representations, and the cognitive operations are transformations of these representations. In this sense, information processing is the act of performing cognitive operations, and the function of brain structures is to perform cognitive operations. We are therefore also arguing that the scientific evidence does not support the proposition that the brain performs cognitive operations.

We wish to consider the experimental evidence that justifies P. We guess there are tens of thousands of papers in scientific journals that may be used as evidence, thus a systematic evaluation of every paper independently

and subsequent integration of the evidence is not feasible. The task must be simplified, but in a way that addresses the initial question. As a first step, we will only consider research that involves measuring or manipulating the physical properties of brain. While a study that does not involve brain properties may contribute to our scientific understanding of the brain, it can only do so indirectly by prompting theory formation and characterizing behavioral phenomena. For example, in 1908 Yerkes and Dodson discovered that performance on a task at first increases with increasing arousal and then decreases once arousal levels become too great. They quantified the intuition that 'stress' can enhance performance. This interesting and useful finding may suggest neurophysiological correlates of performance and arousal, but it does not experimentally justify any statement about brain function in a direct sense.

We are primarily interested with research that investigates the relations between physical brain properties and behaviors, abilities, or physical (sensory) contexts. Cognitive scientists, and philosophers of science, typically reference the evidence from this category of research when making claims about brain function. Relational evidence, as we will call it, can be broken down into four major categories:

1) *structure/ability* studies,
2) *external-stimulus/brain-response* studies,
3) *task/brain-response* studies, and
4) *brain-manipulation/behavioral-response* studies.

In structure/ability (or S/A) studies, researchers relate the structure or structural states of the brain to the absence or presence of particular behaviors or abilities. Paul Broca (1861) popularized this type of research with his lesion-deficit, or lesion study, when he discovered an individual who could only speak the syllable 'tan'. A post-mortem analysis of the person's brain revealed damaged brain tissue in the posterior part of the left inferior frontal gyrus—a region now known as Broca's area. Thus Broca related the ability to produce fluent speech to the left inferior frontal gyrus. The class of S/A studies includes more than lesion studies since any physical feature of the brain (e.g. patterns of white matter connectivity) may be associated with the absence or *presence* of specific abilities.

The structure in S/A studies refers to the physical structure of the brain as measured by a variety of measuring techniques, the most basic being gross anatomical observation of brain tissue. Other measuring techniques include, but are not limited to, histological examination, molecular analysis, electroencephalography (EEG), magnetoencephalogray (MEG), computed tomography (CT), positron emissions tomography (PET), single photon emitted computed tomography (SPECT), structural magnetic resonance imaging (MRI), diffusion tensor imaging (DTI), magnetic resonance spectroscopy (MRS), and others. By ability we mean any observable behavior that can be done by an animal or human, such as the ability to count out loud, to raise one's arm, to navigate a maze, to write a sentence, to score above chance on a test, etc.

In external-stimulus/brain-response (ES/BR) studies, the experimenter systematically manipulates a physical feature of an organism's external environment, and measures temporally coincident properties of the organism's brain. The response need not occur precisely simultaneous with the stimulus and is typically extended in time. Edgar Adrian is generally credited with pioneering stimulus-response studies of nervous tissue. He was the first to record the electrical activity of single nerve fibers, and was subsequently awarded a Nobel Prize in 1932 for his work. As an example of his prolific research, he isolated an eel's eye and optic nerve, attached electrodes to the nerve, and recorded the electrical activity to varying lighting situations (Adrian & Matthews 1927).

Brain responses in ES/BR studies are recorded using a variety of techniques based upon electromagnetic brain properties, including single-unit intra and extracellular recording, evoked potentials, EEG, MEG, and others. Functional MRI (fMRI) is a popular tool used by cognitive scientists to assess brain properties in response to an ES. Proper interpretation of the fMRI *signal* itself requires technical background knowledge (Roskies 2007; Logothetis 2008). Briefly, the fMRI signal is a consequence of the magnetic properties of blood components which covary with local metabolic demands. Other common techniques in ES/BR studies include PET and optical imaging.

When, within a research protocol, a brain-response is recorded while the organism is performing a particular task or activity, we call this a task/brain-response (T/BR) study. The form of the T/BR study is similar to ES/BR studies, except the process of completing the task is 'self-directed' rather than under complete control of the experimenter. Often T/BR studies include ES aspects as well. Memory research provides typical examples. Poppenk et al. (2010), in studying prospective memory which is described as the ability to act out postponed intentions at future times, presented a series of visual scenes (ES) to subjects and instructed the subjects to either imagine performing an action associated with the visual scene, or to use the scene as a reminder to perform an action the next time the same scene was viewed. FMRI was used to measure properties of the subjects' brains during the tasks. Notice that although the experimenter controls the ES and the task *command* directly, she cannot control the *process* by which the subject completes the task.

Brain-manipulation/behavioral response (BM/BR) studies differ from ES/BR and T/BR studies in that physical properties of the brain are directly controlled or manipulated while a behavioral response is observed. Technologies used for BM include lesioning, gene-expression, direct current stimulation, electrode-based deep brain stimulation, transcranial magnetic stimulation, and light-based optogenetic stimulation, among others. Optogenetic studies are a relatively recent advance. They are based upon the introduction of light-activated channels into specific populations of neurons, permitting relatively precise control of action potential generation in live organisms (Zhang et al. 2007). For example, Wyart et al. (2009) expressed light sensitive genes within so-called Kolmer-Agdur cells of the zebrafish, and then non-invasively manipulated the neuronal activity of these cells which modulated the swimming behavior of the animal.

## 3. Interpretations of the evidence

We wish to determine if research studies from the categories of relational evidence discussed thus far justify the scientific claim P that the brain processes information. Again, we will not consider every study; rather, we will start with a typical example from each category of evidence and attempt to generalize our conclusions to the category.

### 3.1 S/A studies

Beginning with Broca, what can we infer from his S/A study of subject 'Tan', who appeared to understand speech but could only speak the syllable 'tan', and more specifically, how do we clarify the relationship between a brain structure and the inability to produce fluent, complex language (Broca's aphasia)? A few preliminary remarks are necessary. We realize that the scientific community's understanding of Broca's aphasia has grown tremendously since the time of Broca, and that our simplistic description of Broca's aphasia as the inability to produce grammatically correct language is a gross description that, although clinically standard, has been challenged by experts and is surely incomplete (Grodzinsky 2000). Here we are focused on the process of relating brain structure to function. For our purposes an accurate description of Broca's aphasia is unnecessary because our argument will challenge the logic of inferring brain functions given correlations between brain structures and observed abilities in general.

It is clear that Broca's area lesions and Broca's aphasia are related in the sense that they can occur contemporaneously within a single individual. It is also clear that Broca's lesion and Broca's aphasia need not be contemporaneous, for 85% of patients with chronic Broca's aphasia have lesions in Broca's area, and only 50-60% of patients with lesions in Broca's area have a persisting Broca's aphasia (Dronkers 2000). Further, surgical excision of Broca's area—a brain manipulation/behavioral response study—has led only to transient mutism followed by recovery of the patient to normal. This evidence alone is enough to at least challenge claims of understanding the function of Broca's area. One can of course speculate on the functioning of a brain area given an imperfect statistical correlation between that brain area and an observed ability, but any claims of knowledge of function are excessive: we have knowledge of the correlation and not the function.

Let us assume, falsely, that Broca's lesions and Broca's aphasia are perfectly correlated in the sense that 100% of patients with Broca's aphasia have lesions and 100% of

patients with lesions have aphasia, for it is possible that other S/A studies exhibit perfect correlation. From this finding, can we conclude that perfect correlation between structural brain states and particular abilities justifies P? This would at first appear to depend upon the nature of the ability being studied, but it is not clear that any S/A study could empirically justify that the brain processes information. In S/A studies, one may try to infer brain operations given a correlated ability. There is a tendency to argue that the structural brain area that is correlated with an ability is in fact performing a process proximately responsible for the ability, but this is not the logic of S/A studies. The logic of S/A studies is as follows: if a subject with structural pattern S cannot do A, but a subject without S can do A, then the ability to perform A must *depend* upon S in some way. Even if we accept this logic—which requires that the subjects are similar in every other way except S—we cannot logically infer that the *function* of S is to perform an operation 'directly responsible' for A. With regard to Broca's aphasia, the fact that a person produces agrammatical speech after a brain lesion does not logically imply that the damaged brain area organizes linguistic grammar.

It should be clear that directly attributing function based upon S/A studies is not logically justified. Consider this example. There were many times when my computer, the computer I am using to write this, loses a particular ability that I expect it to have. I recall a time when I was unable to run programs I typically run, and other programs began running very slowly or would shut down for no apparent reason in mid-session. The problem turned out to be a dead CPU fan. Should we say that the function of the CPU fan is to run programs quickly and prevent them from shutting down? The abilities in question correlated with the spinning of the fan, but the fan does not perform the absent abilities—the function of the fan is to cool down the CPU. The CPU fan participates in a series of causal interactions that run programs when 'everything is working', and we could similarly argue, given the results of a S/A study, that a structural feature of the brain participates in causal interactions that realize a particular ability under certain conditions. This does not entail, given a S/A study, that the function of the brain structure is to perform a process proximately responsible for the ability in question.

In S/A studies, functions are always speculatively inferred from observed abilities. We do not study how the dynamic processes of the brain area and its relations to the rest of the organism causally make the ability possible—but this is presumably what we need to understand if we are to assign a function (or operation) to the brain area in question. Broca's area is more *selectively* related to the ability to produce grammatically appropriate speech than some other parts of the organism. One may argue that the selective correlation between Broca's area and Broca's aphasia justifies the scientific claim that Broca's area processes linguistic information, but the ability 'to process linguistic information' plays no obvious role in the study. To justify the claim that Broca's area processes linguistic information given the evidence of a related S/A study—a claim made but many scientists in the field—we would have to *assume* that speaking grammatically appropriate speech involves linguistic information processing by the organism somewhere (because this is the only way that processes and operations enter into S/A studies. We do not observe brain area operations; we observe the functioning organism), and then further identify the ability to process linguistic information with the function of Broca's area. But neither of these steps is empirically justified. No one directly measures linguistic information processing in the study—we simply observe the form and content of speech—and the identification of an organism's ability with the functioning of a brain area is a speculative inference. This does not imply that Broca's area plays no role in the ability to produce fluent speech—we simply do not know what that role is given limited evidence and theory.

There is another problem with inferring functions from S/A studies which follows from the theory-ladeness of observations. In S/A studies, we observe abilities and physical brain properties. While there is less debate about how to describe brain properties, the way we describe abilities is not well-defined. This presents a problem, for the function that we attribute to the brain area becomes dependent upon how we describe and interpret the ability that is absent. With regard to Broca's aphasia, Broca himself did not describe the absent ability as a deficit in grammar or aphasia, but rather as an articulation deficit, focusing on the motor aspects of speech. Seen this way, we would infer that the function

of Broca's area is to orchestrate the motor production of speech, or process speech motor information. Later commentators have sought to increasingly refine and abstract their interpretations of the absent ability in Broca's patients. For instance, Grodzinsky (2000) 'observes' that the absent ability demonstrated by Broca's aphasiacs is a specific syntactic ability:

> …the computation of the relation between transformationally moved phrasal constituents and their extraction sites (in line with the Trace-Deletion Hypothesis)...the construction of higher parts of the syntactic tree in speech production. (p. 1)

Thus Grodzinsky says that the function of Broca's area is to perform this very specific syntactic ability, but his interpretation is based upon a relatively advanced linguistic theory *that is not a fact but a theory itself*. As our theories of language and language producing organisms advance—as they did through Chomsky's work, for instance—our descriptions of the absent abilities may change, throwing suspect upon seemingly empirical statements of brain function. Perhaps this criticism is too general, but our way of interpreting abilities is not as firmly grounded as our measurements of brain properties, yet the cognitive scientist projects this interpreted ability onto the brain structure, claiming that the interpreted ability is the objective or natural function of the brain structure. Further, we may reasonably describe abilities in mutually compatible ways, all of which are true, but which generate disparate functional assignments for a particular brain structure (which is one reason why cognitive scientists can disagree so strongly about the function of a brain structure while considering the same body of experimental evidence).

Apart from the interpretational problems associated with identified abilities, there is also the problem of identifying which abilities are absent or present in S/A studies in the first place. For example, in Broca's lesions, the inability to produce grammatically correct language may be a salient and obvious deficit that we notice immediately, but this salience may detract from other deficits (or new abilities) that also accompany Broca's lesions. While researchers have subsequently tested Broca's aphasiacs for other deficits—such as mathematical abilities, language comprehension, and others—they have not been tested for every ability imaginable, or even for a broad range of non-linguistic abilities. But certainly Broca's aphasiacs may lose or acquire other abilities that we simply have not empirically assessed. These abilities need not be obvious, and may only become apparent through very specific and creative testing by the scientist. Therefore the operations or functions we assign to brain structures, based upon S/A studies, are dependent upon the small sub-set of organism-level abilities that we have decided to look at; abilities that may be interpreted in different but compatible ways. These problems do not similarly arise when studying physical processes, such as cellular metabolism, where we select to look at particular physical objects and their interactions, and discover the role those objects and interactions play in the organism. For instance, one can study cellular processes (in a gel or Petri dish) independent of the organism and learn much about function. In contrast, we cannot remove a piece of brain tissue from the organism and expect to learn anything about the *type* of cognitive operations it performs.

We will not address other S/A studies as the form of our argument applies to all studies that attempt to relate brain structures and organism abilities. In summary, the evidence in S/A studies is composed of correlations between brain structures and organism-level abilities. To arrive at claims of information processing in the brain from S/A studies, one must first infer that the brain structure performs an operation or function that directly enacts the organism-level ability in question, but this inference is speculative, biased, and not logically sound. In a second layer of interpretation, one must identify the function with information processing, but this identification enters as an assumption independent of the evidence.

### 3.2 ES/BR studies

External-stimulus/brain-response studies address questions of brain function more directly than S/A studies and are probably the largest category of relational evidence. Edgar Adrian, in his pioneering ES/BR research, measured the electrical responses of single sensory cells, such as stretch receptors, while they were fixed to particular weights. He observed that a cell's electrical responses are in the form of stereotyped action

potentials, or spikes, and that the rate of producing spikes increases as the weight increases. Thus the rate, or frequency, of spikes during a fixed time period is able to predict the magnitude of the stimulus.

These early experiments established that single cell responses and stimulus magnitudes may reliably covary with each other. While magnitudes and intensities are important properties of stimuli, they are not the only properties of environmental stimuli that are relevant to an organism. In general, a stimulus may be characterized by multiple properties. For example, an auditory stimulus may be described by its intensity, frequency spectrum, temporal envelope, source direction, source distance, and so on. It is possible that a particular cell responds to one of these properties and not the others, or to some combination of properties, which suggests that a cell may be *selective* for specific properties or features of the stimulus.

Barlow (1953) was perhaps the first to clearly demonstrate the feature selectivity of sensory cells (Reike et al. 1999). By recording the electrical activity of retinal ganglion cells in the frog, he was able to show that the cell's activity covaries with the location and size of a circular spot light on the retina. After systematically varying the light spot's size and location, Barlow determined that the cell's receptive field—the collection of stimulus properties that maximally activated the cell—is a circularly symmetric form called a center-surround field. Spots of light within a small region of the retina activate the cell, but spots of light away from that region inhibit it.

Hubel and Wiesel (1962) greatly extended Barlow's work and discovered cells of the striate (visual) cortex that have surprisingly complicated receptive fields. Two of these cell types are the so-called simple and complex cells, which respond maximally to appropriately oriented bars or slits of light. Some of the cells are relatively insensitive to the location of the bar, while others only appreciably respond to moving bars. In describing these cells, Hubel says:

> We feel that we have at least some understanding of a cell if we can say that its duty is to take care of a 1 degree by 1 degree region of retina, 6 degrees to the left of the fovea and 4 degrees above it, and to fire whenever a light line on a dark background appears, provided it is inclined at about 45 degrees. (Hubel 1962, p. 168)

The evidence from these pioneering ES/BR electrophysiological studies cannot be interpreted without the concept of selective response. Selective response means, loosely, that the cell fires action potentials only when the 'right' stimulus is present. Put more rigorously, selective response refers to two characteristics of neuronal cells: (1) the rate or pattern of firing action potentials (the spike train) covaries with specific stimulus properties, and (2) different cells may respond differently to the same stimulus. Both characteristics are typically implied when referring to the selectivity of cells in ES/BR studies. If someone discovered a neuron that exhibited (1), but on subsequent research discovered that all neurons exhibited (1) in the same way, one would not say that the initial neuron was selective for the stimulus, even though it exhibited selectivity for some stimuli among others. As well, the fact that different neurons respond differently to similar stimuli does not imply (1), since neuronal responses may be random in response to stimuli. Condition (1) is a form of *within* neuron stimulus selectivity, while condition (2) is a form of *between* neuron stimulus selectivity. For ES/BR studies such as Hubel and Wiesel's, when an ES is chosen and controlled by the researcher, we assume that the relation between the ES and BR is causal, as this assumption does not change our interpretation of selectively, even though we use the term 'covaries' which has statistical connotations.

We are now in a position to evaluate whether Hubel and Wiesel's ground-breaking ES/BR studies justify the claim that the brain processes information. In this case we are asking if specific neurons, complex cells of the striate cortex, process or carry information. The experimental evidence consists of recorded responses of complex cells that demonstrate stimulus selectivity in the senses of (1) and (2). It seems that selectivity in the sense of (2) does not provide any justification that complex cells process information; the fact that different cells respond differently to the same stimulus suggests only that the cells are different in some way.

Claims of information processing, if they are justified by this experiment, must follow from the evidence that complex-cell spike trains covary with the properties of visual stimuli, or in causal language, that different visual stimuli cause different complex cell spike trains. Considering the latter causal language, the fact that different causes reliably produce different effects when mediated by the same cell does not appear to justify the claim that the cell processes information, unless one takes that fact to be a definition of information processing itself. Even so, this type of causal relationship appears everywhere one looks. A particular pool ball when hit by other balls with different masses and velocities will undergo different effects. The pool ball may not appreciably move when stimulated by light or sound at typical intensities. The selectivity of the pool ball to acquire different velocities in response to different causal 'stimuli' does not appear fundamentally different than the selectivity of a complex cell, especially if the visual stimulus is taken to be a space-time collection of photons.

On closer analysis, there is a difference between the causality in the pool ball example and the relation between the ES and BR of complex cells. The pool ball example involves direct physical contact and an exchange of energy and momentum, while the causal response of the complex cell is more indirect. Photons travel through the lens of the eye and are absorbed by photoreceptor cells of the retina. Absorption of photons causes the release of the neurotransmitter glutamate at synapses onto so-called bipolar cells, causing the electrical field across the membrane of these cells to become more positive or negative, which respectively increases or decreases the probability of generating an action potential. Bipolar cells have axons that synapse on other cells, and through a series of neuronal connections, influence the membrane potential of complex cells and subsequent action potential generation. The causal chain from photons to complex cell response is complicated and likely includes causal feedback, yet it is not obvious that a complicated causal chain is necessarily information processing.

Even more worrisome is the fact that selective causation need not imply that the BR has any *functional* relation to the ES at all. Nothing rules out the possibility that those selective correlations are accidental—not in the sense that the correlations are statistically spurious, but that those correlations are functionally irrelevant to the stimuli of interest. As an analogy, suppose my computer has a CPU fan with a blue LED light on the fan. The light, however, is unlit and the fan isn't spinning. It happens that when I kick my computer just so on the left side of the front cover, the LED lights up, the fan begins spinning but stops after a second or two, and the light goes out. If I kick it again, just so, it starts up for a second then stops. I can reliably cause the fan to turn on for a bit. When I kick the computer in other places, or shake it up, or sing to it, nothing happens to the fan. The fan is selectively correlated with a specific kick. Perhaps there are hundreds of computers, constructed at the same factory, that behave similarly. This selective, causal relationship does not imply that the fan is functionally relevant to my kicking, or processes kicking information, or represents kicking. This causal relationship may be accidental. Why then, given the evidence of selective responses in ES/BR studies, do many scientists associate information processing with this sort of causation?

### 3.2.1 Justification of information processing from ES/BR studies

There is a strong tendency to associate information processing with the results of ES/BR experiments like Hubel-Weisel's. The spike trains of neurons appear to be relaying specific messages about the external environment to the organism. Claude Shannon (1948), the founder of mathematical communication theory, rigorously defined a model of information transfer that may explain this appearance. In Shannon's language, the physical environment acts as a source that generates a message (ES), the message is transformed by a transmitter—a sensory organ of the organism—into a signal suitable for biological transmission. The spike train (BR) is assumed to be this signal and the neuron to be the transmission channel. These comparisons are reasonable, but the next stage of the communication model, however, is problematic. Communication requires a receiver that performs the inverse operation of the transmitter, or something that reconstructs the environmental message from the spike train signal.

The experimental researcher, the one who discovers selective correlations between neuronal spike trains and environmental messages (stimuli), often plays the

surrogate role of the receiver or decoder. By describing relational or mathematical mappings between the ES and BR, neuroscientists attempt to 'read the neural code.' But this is not the sort of information transmission we were trying to explain. To complete the biological communication model, and to ground information transfer, we need to explain how the organism can reconstruct the environmental message from its temporal pattern of action potentials, and we must demonstrate that the organism reproduces a similar environmental message within the organism itself. The neuronal spike train is not the message—if anything it is the transmission signal or 'encoded message.' Although interesting, it is not enough to show that spike trains have the *capacity* to represent environmental messages through selective covariation. The fact that researchers can mathematically map spike trains back onto stimuli does not say anything about how the organism physically reconstructs the environmental message. This capacity to map follows immediately from statistical correlations. Neuroscientists who acknowledge these limitations explain that mathematically reconstructing stimuli from spike trains requires taking the homunculus point of view (Reike et al. 1999).

For an organism to receive an environmental message in Shannon's sense, that message must be within the organism and have the same structure as the original message. This suggestion may appear radical, but it is simply the completion of Shannon's communication model—the same model that supports the intuition that the brain processes and transmits information. For example, consider telephonic communication. Air pressure waves may be converted into analog electronic messages that are encoded into digital signals and transmitted through a physical channel. This digital signal, which does not mirror the sound wave in form, reaches a destination where it is reconstructed back into an analog message that drives a loudspeaker, reproducing the original pressure wave. If the original message was not reproduced (perhaps imperfectly) at a destination, we could not claim that communication or information transfer took place. A message is communicated if and only if that message is reproduced at the receiver.

If one assumes that the organism receives environmental messages, then in accordance with Shannon's communication model, at least the structure of that message must be physically reproduced within the organism. The alleged *encoded* message—or spike train—has a physical basis, thus the message ought to have a physical basis as well. This means that the scientist would have to demonstrate a set of brain-related physical measurements that copy, perhaps imperfectly, the structure of an environmental stimulus. Let us call this the brain-image of an environmental message. It would remain for the scientist to describe the mechanisms by which neuronal spike trains causally reconstruct the brain-image of a particular environmental message.

When decoding spike trains in practice, the neuroscientist leaves the animal lab and goes to work at the computer. On the computer, spike trains and environmental stimuli are given numerical representations. The creative work involves finding mathematical algorithms and heuristics—let us call these the decoding procedures—that link spike trains to stimuli. When the neuroscientist finds a decoding procedure that works, she claims to have discovered a neural code. The problem is that the neurons themselves have no physical relation to the decoding procedure. The *actual* neuronal spike trains in the living organism do not reconstruct environmental stimuli within the organism using these fabricated decoding procedures, or at least the neuroscientist has no evidence of this. If she supposes that other neurons have the function of performing the decoding procedures that she discovered, and she wishes to find biological evidence, then she must record from neurons that allegedly perform the decode, and, using similar mathematical techniques above, fabricate a secondary decoding procedure that links these spike trains to the original decoding procedures. These investigations lead to an infinite experimental regress that mirrors the epistemological regress of the homunculus argument. The only way to stop the regress is to discover the brain-image of the stimulus.

But no evidence suggests that brain-images exist, so the very presence of an encoded message within the brain presents a problem. In other words, why should the brain contain encoded messages that transmit environmental messages, yet never reproduce the structure of the message itself? The organism requires the actual message, and not only an encoded version of it. At this

point our analogy to Shannon's communication model breaks down. It does not appear that the environment communicates a message to the organism, but rather, the organism is perhaps translating the environment. Spike trains are not signals corresponding to encoded messages; they are the actual messages only in the language of the organism, whatever that might mean. With respect to the organism, the message is not encoded in anyway, and speaking of a neural code is metaphorical and at times misleading. The analogy has changed from information transmission to language translation. But even the idea that spike trains are a language is metaphorical—spikes trains need not constitute a private biological language, and since Wittgenstein, philosopher's have questioned the coherence of a private language altogether. Our goal here, however, is not to support other metaphors, but to show that Shannon's communication model, which is an integral part of modern technology, does not match the relation between an ES and BR.

The decoding procedures discovered by neuroscientists are useful in that they allow us to predict spike trains given environmental stimuli, and stimuli given spike trains; but the specific decoding procedures do not tell us anything about the function of neuronal populations—because the decoding algorithms have nothing to do with the biology of the organism. Rather, the capacity to successfully predict between stimuli and spike trains via decoding is typically taken as evidence that spike trains represent stimuli, although the capacity to predict immediately follows from the statistical correlations between spike trains and stimuli.

There are neuroscientists who consistently, and with clearly stated assumptions, apply Shannon's mathematical information theory to neuronal data with the goal of quantifying the theoretical channel *capacity*, or bit rate, of spike trains (Strong et al. 1998; Reike et al. 1999). These interesting applications of information theory within neuroscience try to answer the following question: assuming spike trains carry Shannon information about the environment, how much information (in bits) could they carry? We could ask similar questions about the oxygen molecules in one's living room, the ants in an anthill, or the blades of grass in one's yard—although the answers presumably would not be as interesting. The fact that Shannon information theory can be rigorously applied to spike trains does not imply that the brain processes information.

Other neuroscientists, such as deCharms and Zador (2000), repeatedly claim that spike trains carry information about the environment as a fact, and suggest what it means to carry information: "Imagine recording from the neuron labeled B1 during different types of stimuli or behaviors and discovering the information that this neuron carries about the organism's environment—the content of this neuron's signal" (p. 614-15). In a concrete example about a retinal cell they say that "The activity of the neuron will be highly correlated with the point of luminance (thus carrying content about this input)" (p. 637). Like in Hubel-Wiesel's ES/BR experiments, we call this evidence the selective covariation between stimulus properties and spike trains. deCharms and Zador use the word 'information' above to possibly mean 'specific properties or features of the stimulus.' Given these examples, we can suppose that they would endorse the following argument: (1) spikes trains and stimulus properties selectively (and causally) covary, and (2) the (representational) content of a spike train is the stimulus property that causes that spike train.

deCharms and Zador do not bring forth any other types of experimental evidence other than selective covariation to justify the claim that spike trains carry informational or representational content, although they do stress that the representational nature of spikes trains is based upon content and function. We have argued that (1) is a statement about the evidence that all of us would agree upon, but that (2) does not obviously follow. The fact that an ES and BR selectively covary, through causal paths, does not appear sufficient to justify claims of representational content, and it has been argued that covariation of this sort is not even necessary for representational content (Millikan 1989; Bechtel 1998).

We need not expect deCharms and Zador, as neuroscientists, to philosophically justify what it means for a spike train to carry informational content, yet if claims of carrying content do not follow immediately from the observed evidence, then we can only assume that they are interpreting the evidence or communicating the evidence by way of metaphor. But deCharms and Zador, along with many other neuroscientists, speak as

though 'carrying content' is a straightforward experimental fact apart from, or in addition to, selective covariation.

From a philosophical perspective, Dretske (1981, 1995) argues that regular causal covariation, by itself, implies information carrying. For example, he says that flag poles and metal paper clips carry information about temperature because the volumes of these metal objects are reliably correlated with temperature. But is it not too easy to find this sort of information carrying all around us? And why do the objects in question need to be regularly or reliably correlated? Any two things that are causally related, perhaps probabilistically, transmit the same sort of thing. If the flag pole was hit by a lightning bolt, does not the flag pole carry information about the energy of the lightning bolt?

So long as the causal relations are understood between objects *c* and *e*, then we might say that *e* carries information about *c*. If a situation can be expressed in the form of a law-like equation, then any parameter on one side of the equation can be said to carry information about a parameter on the other side equation, such as the ideal gas law PV=nRT. If the conditions are probabilistic, then we can use probability theory to derive the distribution of one variable given another, so long as we have some understanding of the physical connections between variables. Carrying information, at least according to Dretske, follows directly from knowing the causal relations between two physical situations, or from minimally knowing that two situations are statistically correlated. If by processing information neuroscientists and philosophers mean that stimulus properties causally or statistically covary with regionally specific neuronal activity, then we agree with P, although we suggest abandoning P in favor of more empirically-grounded statements about covariation.

### 3.2.2 Other philosophical justification of information in ES/BR studies

Considering similar ES/BR experimental evidence, Garson (2003) has attempted to explain a concept of information based upon the pioneering electrophysiological ES/BR studies of Edgar Adrian. Hubel-Wiesel's and Adrian's experiments were similar; both consisted of presenting stimuli while measuring the electrical responses of single cells. Although the technologies, organisms, cell types, and stimulus types differed between Hubel-Wiesel's and Adrian's experiments; the evidence in both consisted of the relations between stimuli and neuronal spike trains, and it is this evidence that Garson uses to elucidate a concept of information.

To ground his concept of information, Garson argues—in accordance with Adrian—that "differences in the frequency of the sequence of action potentials map onto the differences in the intensity of the stimulus that produce them, and not to a constant state of the stimulus." (p. 931). He argues that *differences* in stimuli and spike trains map to or covary with each other, and that this fact captures the sense in which spike trains and stimuli are *arbitrarily* related to each other, thus supporting the informational nature of the spike train.

Garson's argument stands or falls with the truth of differential mapping, yet his belief in differential mapping, which he takes from Adrian's work, does not correspond to the predominant experimental methodology used to acquire evidence in ES/BR studies. Since Hubel-Wiesel's experiments, neuronal responses in electrophysiological ES/BR studies are most often understood by characterizing the feature selectivity of the cell type. This selectively corresponds to the collection of stimulus properties that evoke responses for that cell, and highlights the properties that evoke optimal responses. It is based upon the concept of selective response that we analyzed above. Selectivity typically involves a non-differential mapping between stimuli and responses, it grounds our current understanding of sensory cell types, and directly opposes Garson's concept of differential mapping.

But the fact that a scientific community makes use of feature selectivity rather than differential mapping to acquire evidence does not in itself deny differential mapping. We, too, are questioning the community in its interpretations of evidence, although we are not challenging its methodology. Garson argues that the differential mapping between stimuli and spike rates logically follows from a conjunction of the principle of neuronal rate coding with the principle of adaptation:

> While the principle of rate coding entails that the frequency of the sequence of action potentials is an exponential function of the magnitude of the stimulus, the principle of adaptation entails that upon application of a constant stimulus, the frequency of the sequence of action potentials will diminish, and eventually such outputs will stop being produced. Hence the relation between the sequence and stimulus is differential. (p. 931)

No further derivation is given, which is concerning since each principle taken individually contradicts differential mapping in Garson's sense. Rate coding is a form of non-differential mapping: a specific stimulus intensity directly maps onto a specific frequency of action potentials. How can Garson assume non-differential mapping as a premise to establish differential mapping? And the principle of adaptation is equally troubling; it implies that a negative change in the frequency of action potentials maps onto a constant intensity of the stimulus. But this directly conflicts with Garson's claim that differences in spike rates map onto differences in intensity, and "not to a constant state of the stimulus."

We agree with Garson that differences in stimuli are particularly important to the human organism and other animals, but Garson does not logically establish differential mapping—nor does the scientific evidence primarily support differential mapping—and thus he does not reach the goal of deriving a concept of information from the evidence.

#### 3.2.3 Summary of ES/BR studies

Rigorous experimental neuroscience has demonstrated neurons that selectively respond to a wide range of measurable parameters across the senses. Because selective responses reliably covary with stimulus properties, we can use mathematical tools to predict stimulus properties given spike trains, and predict spike trains given stimulus properties. However, the model of the environment communicating messages via spike trains is not analogous to Shannon's communication model, at least not unless the scientist who records spike trains is included in the model. Nor is there evidence that the neural decoding algorithms proposed by neuroscientists actually take place within the brain, and it does not seem possible to experimentally show that they do using standard ES/BR experiments. Further, the idea that spikes trains carry information is grounded in the experimental evidence of selective covariations, but most commentators conflate or equate carrying information and selective covariation, while others have argued that covariation is not even a necessary condition of carrying information.

### 3.3 T/BR studies

Task/brain-response studies combine aspects of S/A and ES/BR studies, and our critiques of these studies will apply. In addition to systematically manipulating the external environment as in ES/BR studies, T/BR studies add to this manipulation a task for the subject to perform. The task is similar to an ability in a S/A study, except that the task is a transient activity with conditions for fulfillment while an ability is an ongoing capacity to act in a particular way. With regard to scientific research, tasks should have observable or measurable criteria for successful completion, while abilities should have observable or measurable criteria for possession of the ability. We can often study a topic using either task or ability language. For instance, in memory studies, we can assign subjects the task of memorizing a set or numbers, and then ask for those numbers at a later time. The task of remembering and the ability to remember are similar in that the criteria for completion of the task and possession of the ability are equivalent. If one remembers the numbers correctly, one has successfully completed the task and possesses the ability to remember.

Let us recall the Poppenk et al. (2010) T/BR study on prospective memory described above, where subjects were presented a series of visual scenes and instructed to either imagine performing an action associated with that visual scene (e.g. swinging on a swing when shown a swing), or to use the scene as a reminder to perform an action the next time the same scene was viewed. FMRI was used to measure brain properties while the subjects performed these tasks. After this task, the subjects were taken to a quite room and asked to perform an identification test. They were shown visual scenes on a computer and asked to indicate whether each scene was studied as an intention, an action, or not seen during

scanning at all. Researchers recorded correct and incorrect responses. The results of the identification test were statistically correlated with the fMRI data to identify spatiotemporal fMRI activity patterns that predicted correct responses on the identification test. Given the results, Poppenk et al. speculated that some of the identified brain regions enact "processes associated with successful encoding of intentions" (p. 911).

This particular T/BR study is more complicated than the ES/BR studies of Hubel-Wiesel from an interpretational standpoint. Although T/BR studies are not necessarily more complicated than ES/BR studies, the complexity of Poppenk et al.'s study is not atypical for fMRI studies that include cognitive tasks. Like S/A and ES/BR studies described above, the empirical evidence in this T/BR study consists of selective correlations, in this case between successful task completion and properties of brain areas. These selective correlations, like those described in S/A and ES/BR studies, do not logically imply that the function of the identified brain regions is to perform a process directly responsible for the task. Poppenk et al. make no attempt to understand the processes of the identified brain region other than to cautiously say that the processes are *associated* with the task, but the observed form of this association is statistical correlation. Even if this association was selectively causal, we still could not infer that the function of the brain region is directly related to completing the task, for the causal association could be accidental. And even if the function of the brain region involves processes to complete the task, we do not know that completing the task involves processing information of any kind.

### 3.4 BM/BR studies

Brain-manipulation/behavioral response studies demonstrate the behavioral effects of causally manipulating brain properties. As an example, optogenetic studies are a relatively recent advance in BM/BR experimentation, and permit precise manipulation of neuronal activity. These experiments involve expressing light-sensitive genes within specific neurons or populations of neurons in living animals. When the neurons with the expressed genes are exposed to light of a particular wavelength, the activity of the neuron will either increase or decrease, allowing for precise control of the neuron's activity. Presumably the expressed genes do not significantly alter the functioning of the neuron otherwise. For example, Wyart et al. (2009) expressed light sensitive genes within so-called Kolmer-Agdur cells of the zebrafish, and then non-invasively manipulated the neuronal activity of these cells which modulated the swimming behavior of the animal.

BM/BR studies establish causal relationships between the activity of multiple brain areas or between brain area activity and behavioral responses. Canonical examples of BM/BR experiments involve electrical stimulation of brain areas resulting in muscle movements. These types of experiments can be traced back to at least Fritsch and Hitzig (1870) who applied surface electrodes to dog brain and demonstrated that the anatomical location of electrical stimulation selectively covaried with movements in different muscle groups. Neuroscientists name the structure of this covariation a somatotopic organization, and classically explain the relation between cerebral cortex and muscle movements with the following three hypotheses (Graziano et al. 2002): (1) the precentral gyrus, or primary motor cortex, contains an explicit topographic map of the body with the foot on the top of the cerebral hemisphere, the mouth on the bottom, and other parts systematically organized in between; (2) the activity at each point in the map specifies the tension in a small group of muscle fibers; and (3) cortical motor areas are organized in a clear hierarchy of control. These classical hypotheses have been summarized as far back as 1938 by Fulton, although there is a significant body of subsequent evidence that is not compatible with the classical theory (Graziano et al. 2002).

Contemporay BM/BR studies of primate motor function typically involve inserting microelectrodes into the cerebral cortex of an awake animal, and injecting low electrical currents into the cellular network while simultaneously recording the pattern of muscle movements, allowing the researcher to catalog the causal relations between electrical stimulation and these movements. One hopes or assumes that the patterns of muscle movements in response to exogenous stimulation are similar to those caused by endogenous neural activity, although the induced electrical activity is clearly non-

physiologic, complicating any interpretation of the results.

If we consider the brain to be a mechanical mechanism, what do these motor BM/BR studies tell us about the brain? In other words, knowing that anatomical locations of electrical stimulation and patterns of muscle movement covary with each other, what can we say about the operations or functions that occur in those brain areas? We might say that there is a causal propagation of electrical activity, beginning from motor cortical areas through the central nervous system and to spinal motor nerves that enact patterns of muscle contracture. If we electrically stimulate the brain in other areas, the electrical activity does not propagate to spinal motor nerves. The motor cortex therefore acts as a metaphorical gateway or hub of electrical activity from CNS to spinal nerves, where the pathways are at least partially organized with respect to specific muscle movements.

Claims of motor information processing are not needed to describe these results, but presumably arise when one assumes that areas of the motor cortex naturally represent various muscle groups, or that the precentral gyrus contains a map of the body, but the fact that electrical activity propagates through specific pathways does not establish that motor areas naturally represent muscle groups. Just as the idea of a neural code is a sometimes useful metaphor for communicating scientific results, the idea of a topographic map of the body in the precentral gyrus is a useful metaphor for summarizing the data about causal organization with respect to electrical stimulation.

Compare the structure of the patellar reflex in the peripheral nervous system: hitting the patellar tendon stretches the quadriceps muscle which activates sensory receptors that propagate electrical activity through motor neurons that contract the quadriceps. We can describe the propagation of physical changes without any reference to information processing or transmission. Of course, one can use a metaphorical information language to describe the propagation of electrical activity involving the patellar reflex, but this language adds nothing to our physical understanding of the reflex. Electrical activity does not propagate randomly through the brain; its pathways are organized, and scientists attempt to understand this causal organization with respect to an organism's abilities. BM/BR studies are an important tool for understanding this organization, however, knowledge of causal organization does not imply knowledge of the operations that occur within a brain region, nor does it provide evidence of information processing.

## 4. Limitations of our argument

The most obvious concern with our argument is that we have only considered a subset of cognitive and neuroscience research, and within that subset, only a handful of studies. We then attempted to generalize the sorts of functional inferences that are justified given the evidence classes we examined. We justified our approach by claiming that the canonical evidence types we analyzed are typically taken as evidence that the brain processes information, even though the empirical reasoning that leads to this scientific proposition has not been elucidated. It is possible that, upon considering multiple experiments together, and evidence from multiple classes of studies together, one might be better able to justify P. Our analysis was also reductive and non-integrative in that we separated evidence types apart and argued what each evidence-type alone could tell us about brain function. Nonetheless, no one to our knowledge has attempted such an integration to positively justify information processing, presumably because it has not seemed necessary.

We began this investigation without a specific definition of information processing, which may also be taken as a short-coming of our argument since information processing was assumed to be the theme of this work. We counter that our method, instead, was to consider the evidence first, and to determine what understanding of brain information processing was justified by this evidence. Of course our view of the evidence was biased by a priori assumptions, but we purposely attempted to refrain from imposing too much on the concept of information processing at the start.

In our analysis we neglected the many modeling studies in cognitive science. There is no doubt that mathematical modeling will help us understand the functioning of the

brain, but we have not addressed modeling efforts in this work because modeling in cognitive science is underdetermined, and to our knowledge there are no cognitive models that are largely accepted as representing mental facts. Further, our criticisms of neural coding and decoding procedures apply to cognitive models that attempt to relate brain data to environmental stimuli or observed behaviors. A cognitive model that is constrained by brain data is essentially a decoding procedure—a group of heuristics and algorithms—that maps brain data to hypothetical cognitions and vice versa. The model is 'unverifiable' in that it is impossible to show that the model has anything to do with what the organism is, for want of infinite regress. Cognitive models that are not constrained by brain data were ignored because these models could not empirical justify a claim about the brain. We are not, however, suggesting that all modeling in neuroscience is problematic. Models that are restricted to physical properties of neurons, or populations of neurons and other brain matter and organs, help us understand much about the brain.

It may be argued that we misunderstand the methods of cognitive scientists. Cognitive scientists do not attempt to verify the presence of particular cognitive operations in the brain directly, but rather, they typically begin with a priori theories about particular cognitive processes, and use the results of relational studies to choose between these theoretical cognitive processes (Henson 2006). One might call the claim that the brain processes information one of these hypothetical cognitive theories. We have not found empirical support for this cognitive theory, although others may interpret the evidence otherwise. If one assumes that selective correlation and causal organization imply information carrying, then one will see information processing everywhere one looks. In this sense, information processing is not a scientific theory or fact, but a basic principle or metaphor that many people find useful in communicating and interpreting evidence.

It is perhaps possible to interpret all cognitive science evidence to date without affirming that the brain processes information, but this claim may sound absurd to experts in the field. Many cognitive scientists assume that the brain processes information and direct their efforts trying to figure out what specific cognitive operations take place in the brain and where in the brain they take place. The theory that the brain performs cognitive operations makes these practices possible, but it also restricts our search for brain function and our potential for ultimately understanding the brain.

We are unsure what sort of evidence would empirically justify that the brain processes information, which implies, in the end, that we do not know what it means for the brain to process information, or what it means for a brain structure to enact a cognitive operation. When we look at what the evidence justifies, we can conclude that brain activity is selectively correlated with environmental stimuli and that the brain shows causal organization with respect to stimuli and behaviors. These facts are enough to establish the possibility of predicting between brain activity and abilities, and to establish the possibility of controlling organism behaviors and (reported) perceptions by manipulating brain tissue.

How, then, does theorizing add to the science of cognitive science if we deny the presence of cognitive operations in the first place? By focusing upon specific abilities, by re-interpreting abilities, by sub-stratifying behaviors, by decomposing perceptions into properties or categories; the cognitive scientist discovers new 'observables' to correlate with brain activity. Any particular perception (or ability) may be interpreted in numerous ways, limited only by one's creativity. For instance, the image of a face may be decomposed into a set of color points, into a set of 'eigenfaces', into a set of relative positions of facial landmarks, into 3D surface contours, into a set of radial basis functions, etc. In a neuroimaging experiment, the cognitive scientist can choose a particular decomposition and see if brain activity covaries with differences in the parameter values. There is always the tendency to claim that one interpretation or decomposition is the true interpretation in the sense that the true interpretation is the one that the brain 'naturally uses'. It is perhaps underappreciated that brain activity typically covaries under most interpretations, perhaps in different ways, some more useful than others.

## 5. Concluding remarks

Information processing, as a natural biological function of the brain, is either not established by empirical

cognitive neuroscience, or is a metaphor, a folk-psychological concept that carries an air of scientific rigor. Although not an empirically justified brain function, information processing is still a meaningful concept that helps researchers communicate complex scientific findings. For instance, it is much easier to say that the fusiform face area processes facial information than to say that the rate of neuronal firing in the fusiform face area is (relatively) selectively correlated with the external presentation of facial images—but the later is an empirical finding while the former is an over-interpretation of the data. While interpretations of data are an essential aspect of science, claims that are not sufficiently grounded in evidence pose a threat to the scientific enterprise and erode the credibility of the field. Hanson, a cognitive scientist, raises a similar concern about the *particular* functions assigned to brain structures, for when talking about the function of the inferior parietal lobe (IPL), he says that "describing the IPL in some familiar and yet vague psychological terms creates a hopeless muddle of claims and agendas that get fossilized in the journals and training of graduate students" (Bunzl et al. 2010, p 54.). Information processing, as a general function, is one of those claims.

Cognitive science can proceed without assuming that information processing is a natural function of the brain and still retain its status as a science independent of neuroscience. Theorizing and interpretations are important in cognitive science, but the place of this theorizing resides primarily in interpreting abilities, behaviors, and stimuli; and not in hastily assigning functions to brain structures. Put another way, all of the predicting and controlling that cognitive science discovers can and does occur without definitive knowledge of brain functions.

## 6. References


Adrian, E.D. and Zotterman, Y. (1926) The impulses produced by sensory nerve endings. Part II: the response of a single end organ. The Journal of Physiology 61: 151-171.

Adrian, E.D. and Matthews, R. (1927) The action of light on the eye. Part I. The Journal of Physiology 63: 378-414.

Bechtel, W. (2006) Discovering cell mechanisms: The creation of modern cell biology. Cambridge: Cambridge University Press.

Bechtel, W. (2008a) Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience. London: Routledge.

Bechtel, W. (2008b) Mechanisms in cognitive psychology: what are the operations? Philosophy of Science 75: 995-1007.

Bechtel, W. and Richardson, R. C. (2010) Neuroimaging as a tool for functionally decomposing cognitive processes. In: Foundational issues in human brain mapping, ed. S. J. Hanson and M. Bunzl, [pp. 241-262]. Cambridge, MA: MIT Press.

Berker, E.A., Berker A.H., and Smith, A. (1986) Translation of Broca's 1865 report: Localization of speech in the third left frontal convolution. Archives of Neurology 43(10): 1065-1072.

Brewer, W.F. and Lambert, B.L. (2001) The theory-ladenness of observation and the theory-ladenness of the rest of the scientific process. *Philosophy of Science* 68(3): S176-S186.

Broca, P. (1861) Remarks on the seat of the faculty of articulate language, followed by an observation of aphemia. In: Some Papers on the Cerebral Cortex, ed. G. von Bonin, [pp. 49-72]. Springfield, Ill: Charles C. Thomas, 1960.

Bunzl, M., Hanson, S. J., and Poldrack, R. A. (2010) An exchange about localism. In: Foundational issues in human brain mapping, ed. S. J. Hanson and M. Bunzl, [pp. 49-54]. Cambridge, MA: MIT Press.

deCharms, R.C. and Zador, A. (2000) Neural representations and the cortical code. Annual Review of Neuroscience 23: 613-647.

Dretske, F. (1981) Knowledge and the Flow of Information. MIT Press

Dretske, F. (1995) Naturalizing the Mind. MIT Press.

Dronkers, N.F. (2000) The pursuit of brain-language relationships. Brain and Language 71(1): 59-61.

Floridi, L, (Winter 2008 Edition) Semantic conceptions of information, In: The stanford encyclopedia of philosophy, ed. E. N. Zalta. Retrieved December 18, 2010 from http://plato.stanford.edu/archives/win2008/entries/information-semantic/

Fritsch, G., and Hitzig, E. (1870) Ueber die elektrishe Erregarkeit des Grosshirns, translated by G. von Bonin, In: The cerebral cortex, ed. W.W. Nowinski, [pp. 73-96]. Springfield, IL: Thomas, 1960.

Fulton, J.F. (1938) Physiology of the Nervous System. New York: Oxford University Press.

Garson, J. (2003) The introduction of information into neurobiology. Philosophy of Science 70 (5): 926-936.

Graziano, M.S.A., Taylor, C.S.R., Moore, T., and Cooke, D.F. (2002) The cortical control of movement revisited. Neuron 36: 1-20.

Grodzinsky, Y. (2000) The neurology of syntax: language use without Broca's area. Behavioral and Brain Sciences 23(1): 1-71.

Hanson, N. R. (1958) Patterns of Discovery. Cambridge: University of Cambridge Press.

Hardcastle, V.G. and Stewart, C.M. (2002) What do brain data really show? Philosophy of Science 69(3): S72-S82.

Henson, R. (2006) Forward inference using functional neuroimaging: dissociations versus associations. Trends in Cognitive Science 10(2): 64-69.

Hodgkin, A. and Huxley, A. (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. Journal of Physiology 117: 500–544.

Hubel, D. H. and Wiesel, T. N. (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. Journal of Physiology 160: 106-154.

Hubel, D. H. (1962) Transformation of information in the cat's visual system. Proceedings of the International Union of Physiological Sciences 3: 160-169.

Klein, C. (2010) Images are not the evidence in neuroimaging. The British Journal for the Philosophy of Science 61(2): 265-278.

Kuhn, T. S. (1962) The Structure of Scientific Revolutions. Chicago: University of Chicago Press.

Logothetis, N.K., and Brain, A.W. (2004) Interpreting the BOLD signal. Annual Review of Physiology 66: 735-769.

Logothetis, N.K. (2008) What we can do and what we cannot do with fMRI. Nature 453: 869-878.

Millikan, R.G. (1989) Biosemantics. The Journal of Philosophy 86(6): 281-297.

Poldrack, R.A. (2006) Can cognitive processes be inferred from neuroimaging data? Trends in Cognitive Sciences 10(2): 59-63.

Poppenk, J., Moscovitch, M., McIntosh, A.R., Ozcelik, E., and Craik, F.I.M. (2010) Encoding the future: successful processing of intentions engages predictive brain networks. Neuroimage 49: 905-913.

Popper, K.R. (1959) The Logic of Scientific Discovery, London: Hutchinson.

Rieke, F., Warland, D., van Steveninck, R., and Bialek, W. (1999) Spikes: Exploring the Neural Code. Cambridge, MA: MIT Press.

Roskies, A.L. (2007) Are neuroimages like photographs of the brain? Philosophy of Science, 74: 860-872.

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

Strong, S. P., Koberle, R., van Steveninck, R., and Bialek, W. (1998) Entropy and information in neural spike trains. Physical Review Letters 80: 197– 200.

Uttal, W.R. (2001). The new phrenology: The limits of localizing cognitive processes in the brain. Cambridge, MA: MIT Press.

Wyart C., Del Bene, F., Warp, E., Scott, E. K., Trauner, D., Baier, H., and Isacoff, E.Y. (2009) Optogenetic dissection of a behavioral module in the vertebrate spinal cord. Nature 461: 407-10.

Yerkes, R.M., and Dodson, J.D. (1908). The relation of strength of stimulus to rapidity of habit-formation. Journal of Comparative Neurology and Psychology 18: 459-482.

Zhang F., Aravanis, A.M., Adamantidis, A., de Lecea, L., and Deisseroth, K. (2007). Circuit-breakers: optical technologies for probing neural signals and systems. Nature Reviews Neuroscience 8(8): 577–81.