



BIROn - Birkbeck Institutional Research Online

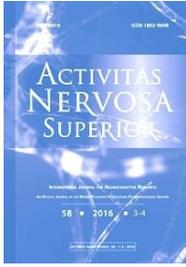
Clapson, Philip (2016) Communication without consciousness: the theory of brain-sign. *Activitas Nervosa Superior* 58 (3-4), pp. 84-107. ISSN 1802-9698.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/18062/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>
contact lib-eprints@bbk.ac.uk.

or alternatively



COMMUNICATION WITHOUT CONSCIOUSNESS: THE THEORY OF BRAIN-SIGN

Philip Clapson*

Department of Psychological Sciences, Birkbeck, University of London, UK

Abstract

Despite developments in neuroscience, consciousness is unidentified in the brain. Moreover there is no scientific definition of what it is or does. This paper proposes that consciousness is not a scientific category. However, by 'postulating' consciousness as self-explanation, the brain can communicate with other brains in collective action. But the brain can generate a more plausible self-description as brain-sign. There are two foundational tenets. (1) Brain-sign arises from the brain's interpretation of its causal orientation towards the world at each moment, and is 'apparent' as the world; and (2) It facilitates communication between brains about the world in collective action which is uncertain or imprecise. It is therefore grounded in the brain's bio-physical operation. Signs are ubiquitous bio-physical states, but they are not causal for the hosting organism. The paper contrasts brain-sign with consciousness both as theory, and in empirical findings. Brain-sign is the source of all theories, including itself.

Key words: *Brain science; Brain-sign; Brain-to-brain communication; Collective action; Consciousness*

1. INTRODUCTION

The many theories developed through the years which propose to explain the brain phenomenon called consciousness are founded on a misconception of its fundamental nature.

Science does not have to "creep up on consciousness", in Andy Clark's words (2016). The supposed ability of consciousness to transcend the organism's bio-physical operation in the world is a prescientific myth. The 'conviction' that we are conscious and have direct access to our mental states is explainable, but it is not a validation. The *brain-sign* alternative (e.g. Clapson 2001; 2014; 2017), proposes how the brain effects inter-neural communication which enables the collective action of organisms in uncertain or imprecise situations. Brain-sign is a wholly physical condition.

(i) The brain-sign model is a description of brain operation, which is applicable to humans and numerous other organisms, and is in line with evolutionary theory.

(ii) It illustrates the boundaries and limits of science arising from the biological nature of human beings. This applies to brain-sign itself.

The paper's structure is influenced by the numerous attempts to reconcile consciousness with the physical brain. Section 2 investigates recent history which has not generated an accepted solution. Section 3 provides an initial outline of the principles of the brain-sign

*Correspondence to: Philip Clapson; e-mail: p.clapson@bbk.ac.uk

Received December 5, Accepted December 28, 2016; *Act Nerv Super* 58(3-4), 84-107; ISSN-1802-9698

model. Other current approaches are compared but, while progressive, they are hampered still by mentalist suppositions. Section 4 provides a more extended overview of brain-sign. Section 5 proposes that the brain-sign model opens a direction for neural science because it escapes the limitations imposed by mind, and provides a practical scientific foundation within the limits of human biology.

2. A BRIEF REVIEW OF THE CONFUSION OVER CONSCIOUSNESS

Given brain-sign's inter-neural role, it might seem to be related to social neuroscience. But social neuroscience suffers from a methodological problem, as it is generally discussed. It arises from the supposition that consciousness does exist and performs a causal role. In the first single authored undergraduate text on the subject, as Jamie Ward claims, he says that: "Social neuroscience can be defined as: an attempt to understand and explain, *using the methods and theories of neuroscience*, how the thoughts, feelings, and behaviours of individuals are influenced by the actual, imagined, or implied presence of others" (2012, p. 10), emphasis in text.

Thoughts and feelings, and other mental states, are facets of consciousness. But Eric Kandel has said uncontroversially in a recent paper: "The unity of consciousness – our sense of self – is the greatest remaining mystery of the brain. As a philosophical concept, consciousness continues to defy consensus" (2013, p. 546).

Ward has no indexed entry for consciousness. Perhaps he judges there is no mystery since he employs the terminology of conscious states throughout the book. But how can the (conscious) thoughts and feelings of an individual be influenced by the presence of others if consciousness is not an accepted, scientifically legitimized (i.e. physical) state? While psychology may be considered by some not a genuine science – one agreed by practitioners in Kuhn's sense of 'normal science' (1996) – neuroscience should be.¹ To explain the influence of others on the thoughts or feelings of an individual would require thoughts and feelings *themselves* to be explainable as neuroscience. There is no extant scientific explanation of mental states as physical and to be found in the brain.

The mind-body problem is not an arcane philosophical topic. It is the fundamental issue in describing human beings and other organisms scientifically, and indeed establishing the nature of science itself as a human (neural) product. Yet discussion has usually involved a fatal limitation: The concept of ourselves as conscious must remain.

Cognitive neuroscience offers no solution. Chapter 14 (pp. 605-643) in the International Student Edition of *The Cognitive Neurosciences* (Gazzaniga et al. 2013), reworked from the 4th Edition Volume edited by Michael Gazzaniga (2009), is titled Consciousness, Free Will, and the Law. Some of this chapter addresses the approaches of various writers: Ned Block, Antonio Damasio, Gazzaniga himself.... However, there is no discussion of a scientific topic, *consciousness*, because no such topic exists. On page 615 is this sentence. "As cognitive

¹For example, in a note in the *American Psychologist*, Kashdan and Steger invoke Thomas Kuhn's book (1996) in this remark. "Kuhn...observed that most of science is spent doing mop-up work, stating that 'no part of the aim of normal science is to call forth new sorts of phenomena' (p. 24). An adherence to safe, publishable science is certainly present in psychology. Certain flaws that proliferate within psychology further compound interference with scientific progress. Specifically, the lack of a uniform lexicon leads to redundant and isolated bodies of work, and in many cases there is insufficient methodological rigor and creativity" (2004, pp. 272-273). Kuhn considered psychology prescientific because it lacked a definite paradigm. For further discussion see, e.g., Hergenhahn and Henley, 2014, particularly pp. 9-11.

The status of psychology as a science has been questioned over its history. Roger Smith (2013) states that the term 'psychology' has been used for very different approaches, two being attempts at a quantitative science which is sceptical about the relevance of mental states, and a more informal human sciences approach (often linked to religion) which emphasises so-called personal experience.

neuroscientists make further attempts to understand the links between conscious and unconscious processing, it becomes clear that these phenomena remain elusive.”

What *is* clear, therefore, is that these writers, cognitive neuroscience generally, indeed all human life, appear to depend upon a physically unknown and undefined property. Gazzaniga and Mangun’s flyer for the next full edition states: “The fifth edition of *The Cognitive Neurosciences* [2014] continues to chart new directions in the study of the biological underpinnings of complex cognition—the relationship between the structural and physiological mechanisms of the nervous system and the psychological reality of the mind.”

But cognitive neuroscience cannot chart new directions in the biological underpinnings of complex cognition because these phenomena remain elusive. It is unreasonable to claim a psychological reality if those making the claim simultaneously admit they have no explicit topic to address. There are many potential factors to consider. Here are three.

2.1. The water of matter and wine of consciousness

The problem of matter and consciousness has been a particular focus of debate in the last seventy years, and centuries before that. But has this debate been soundly based? It presupposes a mental life by which the physical world can be identified. This is the Phenomenological position. In Maurice Merleau-Ponty’s words: “The whole universe of science is built upon the world as directly experienced, and if we want to subject science itself to rigorous scrutiny and arrive at a precise assessment of its meaning and scope, we must begin by reawakening the basic experience of the world of which science is the second-order expression” (1962, p. viii).²

Merleau-Ponty elevates conscious experience beyond the physicality of science. But as what? Later he says: “We do not yet know what to see is and what to think is...” (1968, p. 158). Thus science is evidently a second-order expression of nothing expressible, which is therefore unacceptable as the foundation of science.

As perplexing is this passage from Tononi and Koch (2015, § 2): “That I am conscious, here and now, is the one fact I am absolutely certain of – all the rest is conjecture.” This was Descartes’ view. But Descartes justified his theory by divine collusion, which is not scientifically acceptable. Tononi and Koch continue: “The past two centuries of clinical and laboratory studies have revealed an intimate relationship between the conscious mind and the brain, but the exact nature of this relationship remains elusive.”³ Surely ‘elusiveness’ implies the opposite of ‘certainty’.

Proposed solutions to the mystery (Kandel’s word) are on offer. Here are some often referenced.

A snapshot of Daniel Dennett’s rather fluid expositions is that consciousness emerges from an accumulated hierarchy of competing sub-personal physical states (e.g. 1991; 1996; 2007).⁴ The familiar notion of consciousness suggests there is a viewer of conscious experience (what is seen, felt, thought, etc., by a subject). A Cartesian mistake, Dennett says. Unfortunately we are not then told how or why these accumulated states result in this specific phenomenon.

²Merleau-Ponty involved Phenomenology with the body’s engagement with the world in perception, which influenced many writers, e.g. Andy Clark, Alva Noë and Shaun Gallagher.

³Tononi and Koch propose five axioms for experiential consciousness: intrinsic existence, composition, information, integration and exclusion. Since the first, intrinsic existence, assumes the brain phenomenon is consciousness, their Integrated Information Theory (IIT) is founded on an uninformative circularity. Indeed, they say that “IIT takes no position on the function of experience as such” (§ 5. Everywhere?: c. Consciousness is adaptive). This appears to contradict the claim that “I am absolutely certain”, since the certainty surely results from being conscious (experience) and can only be effective if it functions as such. Yet about the function they posit no position.

⁴The terminology of personal/sub-personal is, or has become a locus of ambiguity.

Indeed, Dennett says there can be no phenomenology. But, as is commented, this is an explaining away rather than an explanation.

Paul Churchland (e.g. 2007), by contrast, proposes that conscious states are reducible to recurrent neural network brain states by the analogy of light and electromagnetic radiation (light *is* electro-magnetic radiation). The problem is that Churchland's theory is a product of his brain, as is the reductive theory of light. So why does the explanation of consciousness and the reductive theory of light, as *knowledge* of the world, depend upon neural states which have no knowledge of anything, being purely physical?

Ned Block (e.g. 1995) takes a pragmatic route by proposing that consciousness can be divided into two types, *access* and *phenomenal*.⁵ The problem is the latter, a subjective state resulting from the 'what it's like' to be conscious, expounded in Thomas Nagel's bat paper (1974). That can interest philosophers, but it need not bother neuroscience because it does not impede the brain's information processing access function, though the two commonly coincide. However, why *subjectivity* should be assumed is not explained beyond Nagel's controversial thesis.

David Chalmers (1995) proposes dualism, which is not helpful since, in conjecturing the 'substance' of consciousness, e.g. qualia, he proposes consciousness has knowledge of its own apparent substance while denying our ability to know consciousness as substantial, i.e. physical. Hence his notion of 'the hard problem'.

Steven Pinker, by contrast, states: "the mind is what the brain does" (1997). Failure to work out how it happens, the so-called hard problem (e.g. 2007), does not, he says, delegitimize its existence.

In the fifth edition of *The Cognitive Neurosciences*, Sid Kouider and Jérôme Sackur state that: "Despite all the progress made recently in the scientific study of consciousness, there are still intense controversies regarding what the theory of consciousness should be. In particular, the lack of consensus concerning the psychological definition of consciousness has rendered the study of the neural basis somewhat inconclusive" (2014, p. 855).

It is obviously inappropriate that "psychological definition" should depend upon the very state which is undefined, viz. consciousness. If consciousness is undefined, what are the perceptions, feelings, thoughts of psychology supposed to be?

A disturbing passage comes from Patricia Churchland. In her recent book, she says this about Bernard Baars' Global Workspace Theory (e.g. 1996): "Wisely, Baars avoided trying to identify the *essence* of consciousness, realizing that essences are an old-fashioned way of thinking about phenomena that impede making actual progress. This contrasts with the approach favored by some philosophers, whereby they try to identify the defining property of consciousness, such as self-referentiality. Baars...us[es] the metaphor of a workspace to convey the rich integration of information from different sources that characterizes conscious states, along with the availability to different functions of information in the workspace. *Workspace* draws on the idea that consciousness is a consumer of products of lower-level processing, that other functions as decision making and planning can access what is in consciousness" (2013, p. 241).

This undermines science as a discipline. Apparently we should not define what we are talking about; then we can make up any story we like. But if, by definition, consciousness is ontologically different from "lower-level" (?) neural states, how do they interact with those states in the workspace? What does "consumer" mean? What does "in consciousness" mean other than a different kind of state from the "lower-level"? These obvious questions have disappeared in a "wise" "avoidance" of "old-fashioned" specification of what consciousness actually is.

It is not merely that these examples are conflicting and confused. Being so many (and more in the literature), consciousness, as any definable *thing*, fails to emerge.

⁵A third category of consciousness is *monitoring consciousness*, which is apparently of less interest.

2.2. Consciousness and the unconscious

Long before Freud, it was recognized that consciousness could not be the sole causal engine of human behavior. The unconscious is now a crucial feature of cognitive (neuro-)science, as Gazzaniga indicates. Here is William Seager: "Scientific psychology both posited and seemed to require unconscious mental processes that were in their essential features very like the more familiar conscious mental processes of perception, inference and cognition. And far from retreating, the contemporary science of psychology, along with the upstart sciences of artificial intelligence and cognitive science, has shown ever more reliance upon the hypothesis of non-conscious mental processes" (1999, p. 9).

However, it is likely that the unconscious (non-conscious, subconscious) is a theory-saving invention to compensate for the indefinability of consciousness itself.⁶ Indeed it makes the situation worse because it secures the assumption that consciousness does exist by mopping up what it does not explain ("these phenomena are elusive").

Consciousness has been associated with light. Because we can see we can know, and we can see because light pervades the world which allows us to see. "I have seen the light" or "It has dawned on me" are light metaphors indicating understanding and reason, as in Plato's Cave parable in the *Republic*. But how mental processes are replicated unconsciously, as Seager and Freud propose, is a mystery. Since "the mind is what the brain does" but cognitive science relies on "non-conscious mental processes", why would the brain generate *any* processes consciously?

This discussion is not novel. But the binding problem of the brain, the generation of a unified and integrated consciousness, likely results from the assumption of this model – i.e., what the brain does is the mind with separate faculties (domains or modularity), notionally in separate neural areas.

2.3. The biological functioning of the brain

To say the brain causes bodily action to fulfil its biological function, to survive and reproduce, is uncontroversial. However, to allow the brain to operate by perceptions, feelings and thoughts is to accept mental states as biological. And since perceptions are purposive, the brain must operate by purposes. (Because *we* see, so we can act for a purpose.) This can be seen in such authors as the mathematician Robert Crease. In his book *The Great Equations*, he writes: "Newton's strange new world was found in our world – but it is not our world, either, nor one we could live in. We humans, even the scientists among us, inhabit what philosophers call the 'lived world', amid designs, desires and purposes: we live in an Aristotelian world" (2008, p. 64).

Modern science eliminates purposes in the physical universe. To suppose we 'live' by them (whatever that means exactly) infects brain science with categories requiring elimination for a physical science to exist.

Crease may not worry about his terminology because brain science is not his topic. Yet Edelman and Tononi write these sentences: "Because a selectional nervous system is not preprogrammed, it requires *value* constraints to develop categorical responses which are adaptive. The diffuse ascending value systems of the brain are known to be richly connected to the concept-forming regions of the brain, notably the frontal and temporal cortex, but also the so-called limbic system...." (2000, p. 105) emphasis in text.

⁶Freud said that "The starting point for this investigation [of psycho-analysis] is provided by a fact without parallel, which defies all explanation or description – the fact of consciousness. Nevertheless, if anyone speaks of consciousness we know immediately and from our most personal experience what is meant by it" (1969, p. 14).

While *value* may be a way of speaking between humans, the brain cannot know anything of value, or use it to make judgments and decisions, for value is an 'Aristotelian' concept.⁷ As to the question of preprogramming, this is reminiscent of these sentences by Antonio Damasio: "Consciousness is just the latest and most sophisticated means of generating adequate responses to an environment...by making way for the creation of novel responses in the sort of environment which an organism has not been designed to match, in terms of automated responses" (1999, p. 304) (slightly reconstructed) which it does by "the world of planning, the world of formulation of scenarios and prediction of outcomes" (ibid., p. 303).

Here is an equivalence between 'preprogrammed' and 'automated'. For these writers there is a way in which thought, judgment and decision arise by a magical transformation of physical neural material into mental states (presumably via the so-called *neural correlates of consciousness*) as which thoughts (planning, scenarios) do their mental work. This is bio-scientifically intractable and fails to address the precise problem. How does the 'Aristotelian' world arise from the physical world?

To sum up. What causes the body's actions? The brain. What causes the brain to cause its actions? Value. Purposes. Motivation. Goals. Emotion. Aim. Fear. Belief. Desire. Memory. Empathy. But, as Dennett might say, these are not explanatory in the physicalist terms required.

And there is a further problem. If we do attempt to describe the brain as a purely physical organ, our communication with each other about the brain (verbal or written) seems to demand the conscious processes, considered unacceptable to the science, to be generated.⁸

3. THE ROUTE TO BRAIN-SIGN THEORY

The prime difficulty with consciousness is the lack of a bio-physical explanation of what it is or does. So why must the brain phenomenon be consciousness? Because consciousness appears to give us *direct knowledge access* to the world, self-evident in our perceptions, feelings and thoughts. But, as the history of science informs us, self-evidence is not necessarily valid.

A definition of seeing is, that the mental subject has direct knowledge of his/her surroundings because they are immediately apparent. But if, as is widely supposed, the transduced impact of electromagnetic radiation on the retina activates already established neural conditions in the brain so that the occurrent state is what Edelman has termed 'the remembered present' (e.g. 2000), i.e. the subject does not see the world but is part of largely preexistent neural conditions, then no mental subject sees the world. Therefore to say we see or perceive (or the brain sees) is neurobiologically misleading.⁹

What tests have been carried out for consciousness as a scientific theory? The answer is essentially negative. Since it is undefined, or elusive, how would tests be construed?

3.1. The brain phenomenon

The use of the term *phenomenon* can be traced to Kant who distinguished the content of the mind, *phenomena* (or *appearances*), from *noumena* (or *things in themselves*, i.e. the physical

⁷The term 'value' is widely used. For example Luiz Pessoa, in a section titled Value Representation, states that "Linking stimuli (or more generally, object representation) to estimates of biological *value* is another important dimension of amygdala function" (2013, p. 26), emphasis in text. The notion of value is consistently linked to reward, again projecting mentalism into the brain.

⁸This is the problem with Paul Churchland's excursion into eliminativism. (See Clapson 2014).

⁹Andy Clark commences his 2016 book by making a similar statement (pp. 13-14). His and others' positions will be addressed in due course.

world).¹⁰ Phenomena are not part of the physical world; they are of the mind, and deliberately so to retain the link with God.

The modern version of this ambiguity replaces Descartes' divinity by subjectivity (a problem Kant also had to deal with). Block's phenomenal consciousness and Chalmers' hard problem are the nature of subjective experience (the 'what it's like').¹¹ But could a physical object, the brain, have any kind of experiences? By putting it this way, these writers are prisoners of philosophical history and its divinity associated narratives.

'Phenomenon', in ordinary language, is used of a purely physical state meaning some-thing which is apparent. The changing cell states of the chameleon's skin are just such a physical phenomenon, though we suppose we know of it by seeing it. However, the brain phenomenon itself can be some-thing. What is necessary scientifically, therefore, is to identify what that thing is. What is mandatory (which distinguishes it from Block and Chalmers) is that the brain phenomenon (the thing) is a state of current physical science, in fact biology, which consciousness is not, despite John Searle's assertions (e.g. 1998; 2014; see Clapson 2014).¹² Science must not only rid itself of the 'what it's like', but the notion of seeing itself – or any kind of mental state. An appropriate narrative of science can then be developed.

Eliminating consciousness does not mean the brain phenomenon is epiphenomenal (as with Thomas Huxley (1874), Ray Jackendoff (1987)), or need not be considered (Behaviorism), or can be ignored by neural science (e.g. Andy Clark's (2013) concern about predictive coding). The brain phenomenon will be identified as *functionally necessary* and explicable.

However, removing consciousness from neural function impacts the explanatory process. Hypothetically, *understanding* the meaning of words/sentences gives verbal knowledge of the world, and both Dennett and Churchland retain this characteristic of mental states in what they write.¹³ But before arriving at the brain-sign explanation of how language works, this text also appears to assume meaning and understanding. The reader is asked to accept it for now; all will 'become clear' when writing and speaking, hearing and reading are 'explained'.

3.2. The significance of social interaction

Sociality is now regarded as fundamentally important both in organism behavior and in brain development; but it is generally considered mind-dependent (as Ward above, 2012). Michael Tomasello has written extensively about the social behavior of humans and other creatures. Here is a representative passage from a recent paper: "Overall, then, we may say that humans coordinate their decisions in collaborative situations, especially by communication, in ways that great apes do not. Once they have done so, they are committed to follow through until everyone gets their just desserts, again in a way that great apes are not. And cognitively, [young children] seem to understand the collaborative activity as a dual-level structure of jointness (joint goal and attention) and individuality (individual roles and perspectives) in ways that pre-figure the organization of many complex human institutional structures" (2014, p. 189).

However Tomasello provides no neural explanation of significant terms: 'coordinate decisions'; 'communication'; 'cognitively'; 'understand'; 'collaborative activity'; 'jointness

¹⁰ Kant derived the usage from Leibnitz.

¹¹Block's *access consciousness* derives, conceptually, from functionalism and the computer model.

¹²The 'solutions' outlined briefly (Section 2.1) attempt to substitute Descartes' God collusion. Dennett and Churchland dissolve the mental-ness of mental states. Block retains the mystery (phenomenality), but maintains functionality (access). Chalmers proposes the problem unresolvable unless a panpsychism of experience pervades the universe along with matter (1995-7): Tononi and Koch concur. This inverts God's collusion by consciousness's universal presence. Pinker lets the problem be.

¹³Specifically, Dennett (1991, pp. 438-440) attempts to overturn John Searle's (e.g. 1992) Chinese Room thought experiment against AI's possibility of understanding by proposing that an adequately complex computer system or robot *would* understand Chinese as a human being would. But he concludes that his account is more art than science – and thus leaves understanding in place.

(joint goal and attention); ‘individuality’; ‘perspectives’. This assumes no need for further explanation. The terms appeal partly to physical interactivity (‘coordinate decisions’ between the ‘individuals’), and partly to mental capacities which facilitate the activity (‘decision’, ‘cognitive’, ‘understand’, ‘attention’). The narrative is so familiar it is to be taken for granted. But science is inhibited by the mystery of these terms. What, as the brain, does ‘cognitive’ entail? Or ‘understand’ or ‘attention’?

In his 2014 book *After Phrenology*, Michael Anderson considers the nature of intrapersonal coordination in wielding a hammer and striking a chisel. He says: “Naturally, the coordination involves more than just muscles and joints but also perception and attention, and the action as a whole crucially involves neural coordination – not as a top-down control process but as part and parcel of the whole-body synergy” (p. 275).

Concerning communication between two individuals, he quotes Dale et al.: “[T]wo people interacting in a joint task come to form their behaviors through compensatory complementary behaviors. These behaviors influence one another locally and incrementally, making the whole conversational performance itself a kind of self-organizing synergy (Dale et al. 2013, p. 56)” (ibid.).

Anderson states specifically that the process is not “top-down control” because he rejects the Cartesian mind controlling bodily activity. Neural activity is part of body function and is co-significant in the overall synergetic activity. He quotes Dale et al. further in saying: “We know through extensive explorations...that speech, gesture and other features of interaction will exhibit coordination.... This can be identified as clusters in the network that become tightly entrained over time (e.g., see Louwerse et al. 2009). As an interaction changes across time, the network structure might change, but the degrees of freedom may stay the same.... (Dale et al. 2013, p. 81)” (p. 276).

By contrast with Tomasello, neural activity is incorporated into the description. However Anderson also refers to perception and attention for which he offers no explanation in neural terms. He does not deal with consciousness, and deliberately so (personal communication, 2015). If the brain network is to be considered part of body synergy (intra- or inter-personal), how is it related to mental categories? He offers no kind of answer.

3.3. Initial description of the brain-sign model

You and I are to have coffee. The pot is sitting on the table. You have a cup in your hand. I pick up the pot and pour. What could be more familiar? But how is this accomplished? Pouring could take place in innumerable ways in space.

We could say our brains effect the action by their entrained physical/causal properties, as Dale et al. propose (entrained = linked). Or we may say that we both perceive and pay attention to the situation and act accordingly (in Anderson’s terms). It cannot be both. If it is the causal properties of the brain, what need for causal properties of perception and attention? On the other hand it cannot be by attention and perception because they have not been identified as causal properties. And we cannot say it is both because that results in causal over-determination – two different activities causing one result, deemed scientifically unsatisfactory.

3.3.1. The foundation of the model

The problem is resolved by brain-sign theory. It derives from the question: How do brains mutually identify the world of their cooperative action? If the causal properties of each organism are purely physical (a vast assembly of neural conditions which are not identical to each other), how do they, as apparently entirely separate from each other, identify *the world* of their jointness in collective action?

Mentalism proposes the actual world is mentally represented, and by the causal characteristics of (e.g.) perceptions, feelings, thoughts, that world can be the domain of action.

The supposition is that *primacy* in effective action arises from the knowledge of the actual world that conscious states have. But of course, whether individual brains do represent the world equivalently, i.e. to what degree the represented world is the same world, leads to unresolved issues of subjectivity versus objectivity. And there is no solution to the puzzle of how conscious representations could have causal properties, “the psychological reality of the mind”.

By contrast the brain-sign account proposes that each brain does represent the world, but the representation is a neural sign, a moment by moment account of the world-goal of the action. It derives directly from the *causal orientation* of each brain, which can happen because each brain’s causality is sufficiently detailed towards the world-goal to re-present it – for example the coffee pot and cup. Because both brains are causally orientated to the coffee exchange, each brain can interpret its own causality to represent what it is orientated towards.

The term *causal orientation* is of primary importance in neural science. The organism’s relation to the objects and events of the world is founded entirely upon possible or actual action, not knowledge. As life progresses, the brain’s fabric develops adaptively as causal programs (*neural assemblies*) which, moment by moment, are in particular orientations toward the world to achieve life sustaining bodily actions (if they work).¹⁴

For collective action, the represented worlds need not be identical (or objective); nonetheless they do fulfill the organisms’ biological requirement by *mutually referencing* the actual world toward which each causality is directed. Why is this biologically beneficial? Because it facilitates extensive degrees of freedom for joint action in conditions of imprecision or uncertainty, as in coffee pouring. Surely the organisms are physically separate, but their action in the jointly signified world does not depend on welding one brain to the other by precise links. This contrasts with the collective action of *Eusocial* organisms which depend upon signals which are physically binding and therefore severely limiting (“pre-programmed”); for example the bee dance in which specific action is triggered by the dancer to the responder from inherited and mimetically developed forms of signing behavior. For example, the responder can be bio-physically *instructed* to seek nectar.

3.3.2. The unit of the communicated

Brain-sign’s reference to the actual world has developed evolutionarily and facilitates complex inter-organism communication. The philosophical notion of intentionality or aboutness, the supposed directedness of mental states upon the world which empowers cognitive activity, is thus disengaged: there need be no mental states. Magic is eschewed.

The ‘images’ of coffee pot and cup (the brain phenomenon)¹⁵ enable one causal operation, which brain-sign theory terms *the unit of the communicated*. Coffee pouring is not the cooperation of two independent organisms, but rather the interaction of geographically separate but causally linked organisms. Linkage depends upon the intervening physical conditions: electromagnetic radiation, compression waves and molecular transmission. Electromagnetic radiation streaming from the pot/cup determines the physical states of each brain (given they are specifically engaged in the process of coffee exchange), and electromagnetic radiation from each body to the other locates their interdependence in the

¹⁴Dick Swaab’s book *We Are Our Brains* begins with these words. “Everything we think, do, and refrain from doing is determined by the brain” (2014, p. 3). This might be considered a truism, but immediately we are on an impossible path. While brains may be considered a physical fact, does thinking exist, or joys or sorrows? To say that no one has seen a thought, a joy or a sorrow, as is often said, sends us directly into a theoretical and practical fog.

¹⁵The use of inverted commas around mentalist terms in the text (e.g. 'images') is to identify brain-sign substitution of those terms for brain-sign function. What that function is will be explained. Eventually all brain-sign terms will have their own terminology, as brain-sign replaces consciousness.

cooperative process. The causal intermediary, electromagnetic radiation, cannot be seen by each individual, not because it is unseeable, but because seeing does not occur.

Thus the attribution of *seeing* to occipital areas as a means for the organism to decide what to do, is unnecessary. Those areas (beginning with the so-called primary visual cortex, V1) are crucial in *control* of the organism in space-time via transduced electromagnetic radiation input to the retina. That the world seems laid out before us to act upon as we will (awareness) is a misconstrual of the function of the brain phenomenon. Moment by moment the brain is self-directed to specific action, whether it is carried out or not. That there appears to be an available landscape for action (in so-called peripheral vision) is the moment by moment possibility of an adjacent action, which in the case of coffee pouring is happening continually. It is not mental awareness.

But if this is correct, why have humans been convinced they do see the world – that their state is visual knowledge? Because the ‘conviction’ is part of brain-sign. Were ‘we’ not ‘convinced’, the communicative biology would fail. Hence Tonini and Koch, and Descartes before them: “It is the one fact I am absolutely certain of.” It is not a fact, but a bio-physical artifice. These theorists ‘suppose’, as mental subjects, they are extractable from the biological conditions of their brain/body operation. This theoretical transcendence may serve inter-neural communication, but it is not a scientific account of physical organisms.

Thus the supposed freedom of our conscious states (to see, to make up our minds) is replaced by the wholly deterministic account of brain-sign theory. Every brain-sign, all its content, is inevitably derived from the brain’s causal orientation in space-time. ‘We’ are, as brain-sign, a series of representational states facilitating inter-neural communication by mutual reference to the actual world of ‘our’ joint actions, whether or not any other organism is around. Of course, although brain-sign is a sign, organisms cannot see another’s brain-sign. Brain-sign’s effect is functionally embedded in the action process.

3.4. A clear differentiation

Brain-sign theory occurs in the history of attempts to identify the nature and role of the brain phenomenon. It will be helpful, at this early stage, to compare the theory to other attempts. In this analysis, brain-sign theory will be further explained by showing how its structure and function overcome inadequacies in those attempts.

3.4.1. The timing and function of the brain phenomenon

Brain-sign is a self-description by the brain of its causal orientation *to* the world to effect inter-neural communication *about* the world. The literature, by contrast, is mostly concerned with a *self*-orientated consciousness, problematically related to the actual world. Here are a few examples.

In *On the Freedom of the Will* in 1839, Schopenhauer said that “man does at all times only what he wills, and yet he does this necessarily” (1985, p. 98). “Subjectively...everyone feels that he always does what he wills.... Every natural being...would feel the same, if it *could* feel” (p. 99). Without establishing this in the brain, Schopenhauer emphasizes that “the behavior of every natural being...must be subject to the law of causality in all its severity” (ibid.).

This is paralleled by Daniel Wegner in his *The Illusion of Conscious Will*. He ends with these words: “We usually assume that how things seem is how they are.... Our sense of being a conscious agent who does things comes at a cost of being technically wrong all the time” (2002, p. 342). “Illusory or not, conscious will is the person’s guide to his or her own moral responsibility for action” (p. 341). That the sense of conscious will derives from the brain is assumed.

Wegner finds evidence in the work of Benjamin Libet et al. (e.g. 1983). His experiments indicated that conscious ‘decisions’ for action are pre-established in the brain’s operation (by

the *readiness potential*) before the subject's awareness of them, thus questioning the free will associated with consciousness. The role of consciousness, as a result, has been much debated.

Michael Gazzaniga proposes that "consciousness flows easily and naturally from one moment to the next with a [post hoc] single, unified, and coherent narrative. The psychological unity we experience emerges out of the specialized system called 'the interpreter' which generates explanations about our perceptions, memories, and actions and the relationships among them" (2012, p. 102). Gazzaniga follows Schopenhauer's idea, localizing it in the brain.

By contrast, Patricia Churchland proposes the brain contains a map of the external world. "Processes in this neural organization model events in roughly the same way that features of a map model the features of the environment" (2013, p. 34). Churchland warns that while usually effective, no *me* reads the brain map in analogy to the person reading the external map. "This disanalogy with using a road map is part of what makes understanding the brain so difficult" (*ibid.*).¹⁶

However, in each case the relation of consciousness to the natural being, conscious agent, person, we, is utterly opaque. For if causality does *not* reside in the 'experiential' states (themselves unexplained) of the natural being (etc.), what function do they serve?

(i) Schopenhauer effectively states that consciousness is pointless in the physical universe, while insisting it exists. E.g. "A person can *wish* two opposing actions, but *will* only one of them. Only the act reveals to his self-consciousness which of the two he wills" (1985, p. 17). Here is the muddled and nonscientific notion of Compatibilism.

(ii) Wegner's conscious agent has no possibility of moral choice as the neural 'decision' has already occurred. (He refers to epiphenomenalism.)

(iii) Churchland rightly says the *me* (or I) makes the brain difficult to understand. Kant's seminal account of consciousness is founded upon the I, and despite contradictions by some theorists,¹⁷ it sustains the narrative of 'I see', 'I feel', 'I think'. Psychological and brain experiments often depend upon the subject's account of their supposed mental states. But Churchland's "difficulty" arises from the erroneous presupposition of the subject, and consciousness itself.

(iv) Gazzaniga's model - "[the left hemisphere's] drive to generate hypotheses" (p. 102) - arose from experiments with split-brain patients. But it is undermined by an infinite regress. Gazzaniga's own brain's "drive to generate hypotheses" generated the interpreter model. But since *this* explanation of his brain's generation of the model could have resulted from his brain ("the [unlocated] drive"), it can continue indefinitely: with no identifiable beginning it has no end. He attempts to resolve this by claiming that, functionally, "[The] drive to generate hypotheses is the trigger for human beliefs, which, in turn, constrain our brain" (p. 102). But this does not stop the infinite regress.

(v) Nor does it legitimize the interpreter functionally, for the idea that conscious belief constrains the brain is founded upon the software/hardware analogy. "The interactions [between mental states and the brain] will only be understood with a new vocabulary that captures the fact that two different layers of stuff are interacting in such a way that existing alone animates neither" (p. 107). But computer hardware does not *generate* software, as the brain is supposed to generate consciousness: humans do. And since consciousness is not explained as physical (as a "stuff"), a fatal non-biological dualism results.¹⁸

¹⁶Intended or not, her statement reflects the debate over whether we see the world or a representation of the world. This was raised in Arnaud's criticism of Malebranche's approach (see Gaukroger 1990).

¹⁷For example Hume's denial of the subject, Dennett's "centre of narrative gravity" (1991), Thomas Metzinger's claim that the self does not exist (2003), and various proposals that the I is constructed by social interaction.

¹⁸Gazzaniga supports the notion of the emergence of consciousness from the brain, but that emergence is still couched in the terms of hardware/software, cf. John Doyle personal communication to Gazzaniga (p. 107).

3.4.2. Causality, signification and the brain-sign 'subject'

Two seminal points arise from the previous section. Firstly, Schopenhauer, Libet, Wegner, Gazzaniga, misconstrue the biology. The brain phenomenon occurs after the causal processes are established because its function is to signify the relevant world in inter-neural communication.

Secondly, by contrast with Churchland, brain-sign theory explains the I from the physical world. The 'sense' of 'I' (our sense of our-self) is the brain's representation of its having a causal orientation towards the world. It remains the same, not because it characterizes personal identity, but because the brain is always causally orientated. It is a marker of this bio-physical function (i.e. of brain-sign) in each organism.

Here is a summary so far.

i) Brain-sign is a neural interpretation by the brain of its current causal orientation towards the world in terms of the world (e.g. the coffee pot and cup).

ii) Each brain-sign conjoins the representation of the brain as being causally orientated ('I') with what it is causally orientated towards. The 'sense' of 'I' does not *see* the world or a representation, nor does it cause the actions of the organism in which it occurs.

iii) Indeed, the 'I' cannot know what the brain's causality entails. And indeed, since each causal condition will be unique, one may assume that no brain-sign is identically repeated.

iv) However, organisms could not interact collectively in the actual world without the joint occurrence of brain-sign. Brain-sign establishes the 'common' world.

v) There is no concern about *free* will because there is no mental will.

vi) The word *illusion* is regularly used in texts about consciousness when its function is being questioned (e.g. Norretranders, 1998). But since consciousness is replaced by brain-sign, the word 'illusion' is inapt. Consciousness is simply a failed scientific theory.

However, further explanation is required, for it may be claimed that we are aware, in our experience, that our I exists: what Kant referred to as *apperception*. Therefore to say the 'sense' of 'I' is the representation of the brain's causal orientation does not account for the awareness that we are an I experiencing. Certainly what we seem to be aware of is not the brain's represented causal orientation, but rather our I-ness *per se*. But just as the 'sense' of 'I' is not the foundation of consciousness, so 'we' are not aware of the 'I'.

The supposed awareness is actually a neurobiological *emphasis* that this brain-sign is being communicated by this organism, i.e. a further event in brain-sign function. While we may suppose our I is often unobtrusive in experience (e.g. we just seem to see something), the emphatic 'sense' of the 'I's involvement is an insistence of this organism in its communication. With humans it is often expressed in language, as we shall come to: 'I see', 'I feel', 'I think'. But it may also occur, along with the expressive noises and signs, in animals. The dog's look or bark while waiting for the ball to be thrown, which for us (as brain-sign), generates the 'sense' that they are a conscious self by analogy with ourselves (supposedly). In neither case (unobtrusive or emphatic) can we (as an 'I') see or be aware that our brain is communicating in the communication; for, as the representation of neural causality in brain-sign, 'our I' is embedded in the brain's communicative activity. The difficulty Patricia Churchland raises about the I (or me) is thus voided.

3.4.3. Escaping the consciousness narrative

Brain-sign theory frees neuroscience from the apparent impasse referred to by Clark in an endnote: the hopeful one (2016, at p. 237), that "our neurophysiological understandings, via computational and 'systems level' theorizing, [will connect] to the shape and nature of human experience", thus solving the mind-body problem; and the doubtful one in which our understandings cannot in fact explain "the very existence of experience itself" (note text at p. 325). The impasse results from a profound misunderstanding, for neither position explains why or how consciousness exists. Without that explanation, Clark's word 'experience' simply

drowns out any alternative before it can be given.¹⁹ Signs, however, are biologically ubiquitous and unproblematically physical.

3.4.4. *Brain-sign versus mental faculties*

There is another aspect of the consciousness concept which must be addressed. Adjacent to the quotation in the previous section, Clark also describes human experience as an “immers[ion] in a world of sights, sounds and feelings” (p. 237). This perfectly illustrates what is at issue. By compelling the reader to the mentalist impasse, he imposes the (dualist) realm in which, for example, objects can be seen. This results from the notion of mental faculties: seeing, hearing, sensing or emotion. Consciousness is faculty exemplified.²⁰

Not only is this unexplained (how could the world be seen by a physical brain?), but faculties are also supposed as unconscious. “The upstart sciences of artificial intelligence and cognitive science, has shown ever more reliance upon the hypothesis of non-conscious *mental processes*” (Seager above, emphasis now added). But could we see without noticing, and is this really seeing?

Consider the coffee pot and cup. We do not see mere shapes (hypothetically) but specific objects with a use and in a context. They could be described – my pot pouring coffee into your cup. Does that mean, when we see pot and cup, our seeing involves thinking (another faculty)? Patricia Churchland, referring to the work of Dehaene and Changeux (2011), which involves the technique of masking, says that:

“When the visual signal *dog* is masked (not consciously perceived), only early visual areas (in the back of the brain) show activation. By contrast, when the visual signal is *consciously* seen, the posterior activity has spread to more frontal regions, including parietal, temporal, and prefrontal areas. Dehaene and Changeux refer to this as *global ignition*” (2013, p. 243).

The visual signal is masked by showing the word *dog* and then immediately the letters XXXXX. The result is a non-conscious perception. If *dog* is left for 0.5 seconds, it will be seen consciously.

Does this indicate that global ignition renders (or is) consciousness? Or is it, rather, that the content of the phenomenon expresses the boundary or limit of the brain’s current causal orientation – i.e. brain-sign? For seeing (or perceiving, Churchland uses both terms) the word *dog* involves (hypothetically) not merely the transduced impact of a shape on the retina, but one with ‘content’: the ‘recognition’ of *dog* with the notional ‘meaning’ of dogness, connoted by the mentalist term ‘understanding’.

For obviously the ‘content’ entails access to other parts of the brain which is not, as Churchland suggests, an endorsement of global workspace theory. Brain-sign theory provides a more adequate scientific explanation for it does not demand, as does consciousness, a prior explanation of what *it* is, or how *it* could have causal powers. (The word ‘ignition’ associates with light and its sense of cognitive illumination without, however, any explanation of what ‘illumination’ is, or how it could be causal, or even of the brain.) Certainly the brains of the subjects of the experiments can communicate with other brains (e.g. of the experimenters) by brain-sign.

¹⁹There are commonly two kinds of text concerning the brain in this specific sense. One assumes consciousness. Hope is expressed that some future scientific development will explain its existence, but functionally it is then ignored since there is no available scientific account of it; the other, in which consciousness is hardly mentioned but a supposed neural architecture and functioning is explored in some hypothetical detail – apparently depending upon consciousness for our grasp of the text. These two approaches have been referred to by Anderson. “When neuroscientists start brandishing the ‘c’ word, there are two predictable reactions: increased public interest and attention and increased scientific scrutiny and criticism” (2014, p. 109). The first involve “enthusiastic adherents”, the second “question whether we should continue wasting *our* energy figuring out why the brain appears to be wasting *its* energy” (ibid.). A third approach, however, is brain-sign theory.

²⁰Clark’s position here appears to contradict his Kantian opening in his 2013 paper.

There is more to say concerning faculties. Anderson (2014) summarizes experimental work indicating that mental states/conditions are not situated directly on brain regions, as in Phrenology. They are generated from accumulated neural activity which is not specific to one (supposed) mental state. "A central point of this book is not just that we do not get selectivity in the brain but that *we don't need it. We can stop looking for it*" (2014, p. 141). This kind of conclusion is anticipated by (e.g.) Karl Lashley, as quoted by Philip Lieberman: "Temporal integration is not found exclusively in language; the control of leg movements in insects, the song of birds, the control of trotting and pacing in a gaited horse, the rat running the maze, the architect designing a house, the carpenter sawing a board present problems of sequences of action...each of which similarly requires syntactic organisation (Lashley 1951, p. 113)" (2006, p. 12).

Whilst this surely points a fruitful route for the science of brain organization and function, it makes the notion of faculties even more problematic. For the manifestations which have been taken as isolable states of mind – perceptions, feelings, thinking – might appear to be an admixture of those states. Thus there are no discreet faculties but rather manifestations requiring an entirely different analysis, and therefore different foundational assumptions. Anderson suggests that: "we adopt a framework according to which individual regions of the brain exhibit not ...the implementation of a single mental operation but rather relative functional differentiation – the development of regional functional biases...[, and these are to be understood] in terms of different loadings on a set of common, neuroscientifically relevant psychological (NRP) factors that together capture as much as possible...the observed variance in...behavioral outcomes" (2014, pp. 150-151).

But why, then, should our vocabulary involve psychological terms at all if mental states *per se*, as grounded in the brain, are rejected? Brain-sign states are not resident states of the brain: they occur uniquely with each occurrence of causality.

The analysis applies elsewhere. Anil Seth, in discussing emotions, references James (1894) and Lange [1885] "who argued that emotions arise from perceptions of changes in the body" (2013, § Interoceptive inference and emotion). Indeed, "More than a century after James and Lange, there is now a consensus that emotions are psychological states that encompass behavioural, experiential, and visceral changes" (*ibid.*). He references, among others, Antonio Damasio (1999).

Seth's particular concern, however, is the replacement of these approaches by predictive coding (PC): "The concept of PC overturns classical notions of perception as a largely 'bottom-up' process of evidence accumulation or feature detection, proposing instead that perceptual content is specified by top-down predictive signals that emerge from hierarchically organized generative models of the causes of sensory signals" (*ibid.*, § Predictive inference and perception).

Seth claims that PC is gaining increasing traction in the brain sciences. However, he offers no explanation of how "emotions arise from perceptions of changes in the body", which are then presumably causal upon the functioning of a hypothetical mind, by PC or otherwise. Thus, while there could be a benefit to brain science in the PC approach (as with Anderson's neural reuse position), there is no advance in the account of perception or emotion as states of the brain. Indeed, the experience of the self in Seth's account is apparently: "related to several contemporary frameworks for understanding emotion and self, including prominently those of Damasio (1999) Craig (2009) and Critchley (2004) [which] each...see selfhood as grounded in representations of physiological condition, with the [anterior insula cortex] AIC emphasised in some, but not all" (*ibid.*, § Relation to other models of emotion and insula function).

These frameworks tell us nothing about selfhood as an active psychological property in a physiological condition. Merely saying so and pointing to a brain location must be seen as inadequate scientific theorizing. For there is no evidence that brain imaging locates a mental state, though it may be evidence of a level of causal function.

4. DEVELOPMENT OF THE BRAIN-SIGN MODEL

So far this account has used the reconstruction of visual perception and the subject ('I') as guiding features of brain-sign theory. The account now broadens.

The current rationale of neuroscience could be said to have two principles: (i) Consciousness enables neuroscience because it is knowledge of the world, specifically the brain; (ii) Locating consciousness (and its faculties, etc.) in the brain is therefore a prime aim.

Brain-sign theory's approach to neuroscience is different. Brain-sign is inter-neural communication, as which science itself is possible. Neuroscientific investigation takes place by the causal brain on itself.²¹ What should be primary in neuroscience, therefore, is identifying the brain's causal mechanisms in interaction with the world and its own functioning. For that neuroscience requires an adequate theoretical foundation.

4.1. The three elements of brain-sign

As a neural sign, brain-sign has three categories. These are:

- i) *Categories-of-the-world*;
- ii) *Categories-of-interaction*; and
- iii) *Brain-sign language*.

Categories-of-the-world are how the actual world is signified by brain-sign. They may be categorized as '*the world*' because they are created so by the brain.

Categories-of-interaction signify how the brain/organism reacts to the actual world, resulting from its ongoing encounters with it.

Brain-sign language is the means by which one brain can alter the causal orientation of another brain.

In all cases, signifying states are brain states not correlates. This results from the evolutionary development of brain structure, operation and function. Brain-sign theory is founded upon the biological ubiquity of physical signs: the shape of a stick insect or flat fish; the changing skin surface of the chameleon or cuttlefish; the facial expressions of animals, and their bodily gestures – the dog's wagging tail; bird song, whale song, Vervet monkey calls. There is nothing arbitrary or bizarre in the brain-sign physiological hypothesis. However, because the brain phenomenon is a sign and not knowledge, scientists can never *know* how the brain sustains its signification.

4.2. Categories-of-the-world

The coffee pouring event is an example of categories-of-the-world. As participants we seem to see the coffee pot, the cup, the other person and their actions, the room we are in, and so on. But our brains signify their causal orientation towards aspects of the environment, moment by moment, as directed by their in-world task (program). The program is one, or a combination of some, of the mass of established brain assemblies formed during what is termed life. These develop over time in the interchange of the brain with the world and other organisms. How these assemblies are to be identified in the brain depends upon future research into neural structure and operation. (Brain-sign theory does not accept the notion of life as endorsing a conscious awareness of the world (Clapson 2014).)

Categories-of-the-world also involve the impact of compression waves (hearing, so-called), molecular transmission (smelling, so-called), bodily impacts (sensing or feeling, so-called). In pouring the cup of coffee there will be the 'sound' of conversation, the 'smell' of coffee, and the 'feel' of the pot handle and the saucer. Indeed, that organisms 'suppose' they hear via their

²¹This is noted by some writers. Michael O'Shea, for example, says "Thinking about your brain is itself something of a conundrum because you can only think about your brain with your brain" (2005, p. 1). But the significance is not identified.

ears is a brain-sign construct. Each of these characteristics (ears as supposedly seen; the supposed hearing of conversation) are neural constructs for collective action by establishing a joint world, including our bodies and their capacities. The extent and density of these representational states result from the brain's constant need to move rapidly between activities. The sense that we experience a coherent and detailed environment, and are capable of performing any task therein is, as science, a mistake. The brain's focus is directed in preparation for, or in fact, specific action (cf. Section 3.3.2). The mentalist notion of *attention* is superfluous.²² This renders redundant the notion that we are conscious yet also reliant upon an unconscious.

4.3. Categories-of-interaction

The communicative *response* of the organism to the world, including other organisms, aids communicative interaction. This category eliminates ambiguity amongst emotions, feelings and sensitivities. It is implausible, for example, that matter can be happy, dissatisfied or curious. However, bio-physical matter (the brain) is affected in its capacity for survival, its preparedness for mating, its efficiency in cooperation. These physical states of appropriate organisms in space-time can be signified by neural self-interpretation as brain-sign.²³

Consider the coffee event. It could involve various different categories-of-interaction in the participants. Your sense of so-called pleasure, for example, could result from the activation of neural conditions which generate (i) unproblematic physical engagement, particularly in consumption which is life-sustaining, and (ii) cooperation with another, promoted in infancy, which enhances action range and protection from possible external threats. This example demonstrates the general principle, that categories-of-interaction decompose into strands of fundamental biological function. Brain-sign theory terms the strands *bio-physical drivers*. The brain has neither time nor ability to identify the drivers in immediacy (see also Clapson 2017). But neural science, as a product of the brain, can.

Development of a lexicon for bio-physical drivers, groups of which amalgamate into categories-of-interaction, is of prime importance. Though somewhat amorphous in type and relative intensity, categories-of-interaction will be identified from neural assemblies, as, for

²²Patricia Churchland (2013, pp. 253-255) offers an admitted imprecise account of attention in trying to unite the notion of consciousness with attention. She describes two kinds: that which grabs attention, like a loud noise, and that involved in voluntary action. She refers to Michael Cohen and colleagues (without specific reference), who propose that attention seems *necessary* but not *sufficient* for consciousness. However, the mentalist notion of attention can be eliminated by eliminating consciousness.

Michael Graziano, on the other hand, proposes to eliminate the mind-body problem with the theory that *awareness* is a neural description of *attention*. Attention is causal neural function; awareness “is an informational model [of attention] in the human brain” (2013, p. 79). Being informational it is, hypothetically, not mysterious. Unfortunately this interesting book is neither adequately rigorous nor explanatory. For example, “Attention...is a mechanistic process.... Awareness...is accessible as explicit knowledge” (p. 51). How can there be explicit knowledge without a knowing subject? Not precisely engaged. Or, “Awareness, the model, lacks all the mechanistic details of neurons and competing signals in the brain.... Why would an organism need to know those mechanistic details?” (p. 79). But how do neurons and mechanistic details generate this knowledge? What truly bio-physical function is served? By contrast, brain-sign is a neural sign, not knowledge; it serves neural communication as a physical sign. Graziano proposes that awareness (the ‘attention schema’) also does, but it never loses the transcendent attributes of mentality. Indeed, he proposes it is causal for the host organism though it is not mechanistic. In other words, the mind-body problem remains.

²³Paul Griffiths says that “One of the main aims of psychology is to get behind the behavior of humans and other organisms and discover how this behavior is produced” (1997, p. 234). What is not required in neuroscience is an explanation of emotions. The terminology has been effective in human communication, but fails as brain science.

example, in neural programs of coffee receiving, with the relevant presence of serotonin (amongst other neurotransmitters).

The notion of *primary emotions* in psychological literature has obscured the physical interface of organism and world. Bio-physical drivers establish the response of the organism to the world in physically expressible terms, not requiring reduction of the mental to the physical.²⁴ They probably also indicate the boundary of human capabilities.

How, though, is inter-neural communication effected? An example is pain. In mentalist terms, pain is considered subjective because it occurs to a conscious subject. But brain-sign is neither subjective nor objective, the first because there is no mental subject (Section 3.4.2), the second being impossible. In brain-sign terms, pain is a category-of-the-world, i.e. the brain causally orientated to the physical states as so signified. However, the response to pain is a category-of-interaction. Compare this with Chris Frith: "The mental experience of pain is associated with activity in the anterior cingulate cortex. People who suffer from severe chronic pain were sometimes treated by cutting out this area of the brain (cingulotomy). After surgery these people still felt the pain, but they no longer had an emotional response to it" (2007, p. 150f).

It can be stated less ambiguously in brain-sign terms. As a category-of-the-world, 'pain' is a practically universal possible physical state of brains which no mental I experiences, feels or suffers. The category-of-interaction is the so-called emotional response (expressed by mentalism as feeling, but not in this case).

Communication with another about a category-of-interaction results from facial expression, body language and/or verbal communication. However, another brain does not also generate the 'pain' (certainly not another's 'pain' – how could it?), but rather an equivalent response to its apparent occurrence, the category-of-interaction. "When you see another person in pain, you feel the anxiety, but not the pain itself" (Gazzaniga 2012, p. 162, referencing fMRI studies). So-called empathy can occur, then so-called sympathy.

It has been proposed recently that mental mirroring is involved. The concept has developed from the discovery of mirror neurons in which the perception by one monkey of another's actions causes equivalent neural activation (e.g. Rizzolatti et al., 1996). Brain-sign theory dismisses mental mirroring as non-explanatory, or having any causal role in producing subsequent sympathetic action. Categories-of-interaction are how brains *communicate* as physical properties. Subsequent action will result from the chain of neural causes.

Mental mirroring is discussed in Graziano (2013, pp. 185-195). He states the mentalist problem exactly. "Mirror neurons may be active when you watch someone else's behavior; but do mirror neurons actually *cause* you to understand someone else's behavior?" (p. 188). Brain-sign theory resolves this by eliminating mental understanding as redundant dualism. It also eliminates, de facto, the supposed causal properties of Theory of Mind. Thus the prime bio-physical driver of another's brain response to someone 'in pain' (and thus the provision of supportive action) is *not* empathy with their pain, but the limitation of their biological effectiveness which their category-of-interaction signifies.

4.4. The coincidence of the categories

Both categories-of-the-world and -interaction derive directly from the causal orientation of the brain. Neither causes the other. They coincide because biological communication about immediate causal orientation entails both. Depending upon circumstances, of course, their

²⁴Luiz Pessoa opens his book by considering whether mental definitions are necessary (2013, pp. 3-5). Quoting various sources, he states that "I will not define emotion and cognition" (p. 5). Quoting Keith Outley et al., he says "Before the age of science, definitions used to be all-important for understanding. Now we know that that it is more important to understand deeply how a particular process works (2006, p. 29)" (p. 4). As with Patricia Churchland and consciousness (Section 2.1), this simply obscures the foundations of science.

relative prominence will vary. This account contradicts Bennet and Hacker who say that "Seeing an ordinary table or chair does not evoke *any* emotional or attitudinal response whatsoever in normal circumstances" (2003, p. 40). Their position derives from assuming mental faculties (cf. Section 3.4.4).

4.5. Brain-sign language

The language element of brain-sign both resolves long-term debates about what language is and does, and how it can function in the physical universe. It also explains how the speech of other animals relates to human language. It eliminates the mentalist notions of verbal understanding or meaning, or their supposed significance to being human. In this sense the continuity of other species with humans is affirmed, and evolution, as a theory, endorsed.

4.5.1. Approaches to language

Language is usually deemed to have two properties: (i) Providing the capacity for thought, and (ii) Enabling communication with others. The literature varies in their relative importance, but the capacity to think is prized because, in principle, it offers intelligent, rational command of the world. Its significance is expressed by psychologist Euan Macphail: "Language is an instance of undeniable behavioural complexity (with clear adaptive advantages) from which it is possible to derive an account of the origin - via self-consciousness - of feeling-consciousness" (1998, p. 235).

This is a sequence of proposals. Only via language can self-consciousness arise because self-consciousness is a cognitive structure which depends upon language. Only if there is self-consciousness can there be feeling, because feeling depends upon a self experiencing. Since animals and babies have no language they can have no self, and therefore no feeling. Indeed, they are not conscious.

However, Macphail qualifies this by two statements on the same page. "We should all appreciate that...there is a huge degree of uncertainty attached to any current hypothesis for the origins of consciousness." And, "It is far from clear that the major contemporary accounts of consciousness could be subjected to experimental test." Both conditions remain. Thus while Macphail builds an argument for his conclusions, he is aware that the entire enterprise could be misconceived.

By contrast, Anderson highlights the apparent distance of language from biology in this comment: "Infamously, Chomsky has argued that the human language faculty could not have evolved by natural selection (Chomsky 1988). More recently, Jerry Fodor has gone from arguing that evolution by natural selection cannot explain how thoughts have meaning (1990) to arguing that evolution by natural selection cannot explain the nature of cognition (Fodor 2000) to arguing that evolution by natural selection is simply ill conceived (Fodor 2007). Because of the prevalence of these attitudes in the cognitive sciences generally, and in the cognitive neurosciences particularly, a truly functionalist cognitive science has yet fully to emerge" (2014, p. 296).

Anderson proffers a different approach: "It is worthwhile to consider...that language...as with all artifacts and cultural practices, must itself be adaptable to the brain.... That language is learnable without special, dedicated neural machinery should cease to be particular surprising" (pp. 262-263).

He references work indicating that: "Because our language has [a] serial character, language acquisition can be accomplished using existing neural capacities for sequential learning, themselves likely evolved originally for behavioral control purposes" (p. 262).

This reflects the quotation earlier from Karl Lashley via Lieberman. However, because Anderson has not wholly rejected the mentalist narrative, he confuses a crucial distinction between language "adaptability to the brain" and "behavioral control purposes", which

echoes Gazzaniga's left-brain interpreter constraining the brain by mental means (particularly in: *Language and self-control*, pp. 267-268).

Just as mental states lock the neural account into an implausible dualism, so does the capacity for thinking as a higher control function of the brain. So how do we redefine the supposed occurrence of language taking place in our minds?

4.5.2. *The brain-sign account*

The bio-physical function of language, brain-sign theory proposes, is for one brain to modify the causal orientation of another via the intermediary conditions of compression waves or electromagnetic radiation. (This does not result from one person intending to, or knowing they are going to.) Mutual causal orientation renders the possibility of collective action. But it must be signified. How is it initially constructed?

The infant's brain develops its causal orientations to the world by bodily engagement with the world. The association of the infant's brain with an adult pointing to an object and naming it is not an acquisition of (mind) knowledge, but rather the ability of the infant's brain to associate 'saying' the word (i.e. training the bodily processes of 'speech') with its developed causal orientation to the world. Suppose the word is *dog*. What is the status of the apparent neurobiological occurrence of *dog* in the head of the infant when it supposedly hears *dog*, or says *dog*?

There are two aspects. For *dog* to occur in both infant and adult brains (i) the causal orientation of both brains must be to the dog, so (ii) there is inter-neural communication about the presence of the dog by the equivalent brain-sign of *dog* derived from each causal orientation. The 'sense' of understanding the meaning of *dog* (a category-of-interaction) is the possible or actual communication with other brains in its physical occurrence. This applies to the whole vocabulary and structure of a language.

However, that we suppose minds or persons do understand language is how the brain has falsely self-explained the process as brain-sign for thousands of years. Words are not things/objects, as of the world, functioning causally in the mind. (In particular, experiments purportedly demonstrating that thought in paralyzed individuals can cause limb action via computer signaling connections to the motor cortex is a misconstrual of brain operation (e.g. *The Atlantic* 2016). The individual 'supposes' their limb actions are caused by their thoughts, but the thoughts (so-called) are brain-sign communication from preceding neural causality.)

Of course the word *dog*, once acquired, can occur without a dog present. Most of language operates without the presence of what is 'talked' about. But still, the mechanics in the brain remain the same. If A 'tells' B to come to the table for supper, both world-referencing brain-signs result simultaneously. But the brain-sign does not cause B to come to the table; the cause is the neural orientation resulting from compression wave action on the brain by A, from which B's brain-sign derives. If one 'reads' a manual for getting around the city, it is not the brain-sign words or maps that generate successful action; it is the causal orientation resulting from electromagnetic radiation impact from the eyes to the brain, from which brain-signs derive, and existed in the writer years ago.

Science results from the ability of brains to communicate about *the* world as categories-of-the-world, via brain-sign language.

Human languages which, for non-speakers, are mere 'noise' (or incomprehensible text), indicates that language develops as connections between the mechanics of 'speech' (sign or writing) behavior and neural causal orientation. Brain-sign theory clarifies this account by dispensing with meaning and understanding in all languages.

'Words' do not usually maintain a strict particularity. Since causal networks adapt over time, associated 'speech' behavior changes - e.g. cap (headgear), cap (bottle closure), cap (admit to membership of a team), etc. In this they develop bio-physically (contra Fodor).

Other creatures communicate with each other by similar principles. The difference lies in the human brain's capacity for "behavioural complexity". However, in a limited way humans can respond to the dog's utterances, as the dog can to the humans'. Neither hear sounds.

4.5.3. *Introspection*

John Locke's notion that we can look into our mind or introspect is, as brain-sign, a brain model of its operation; but it is incorrect. The brain models looking into something inside itself as it models what 'we do' in the actual world. In 'trying to remember a name', for example, there is a 'sense' of looking inwards and a place to look – the mind. The model has the 'I' as the brain's causal orientation, and an inside mind, currently a blankness, as the category-of-the-world store of information. (Heidegger referred to it ironically as "the 'cabinet' of consciousness" – 1962, p. 89). When the word appears – 'I remember it!' – it is generated by the brain as a new brain-sign, 'I' and 'word'.

But it was not found by our looking. There is nothing for the brain to look into, and no one to look. But since brain-sign is the means of inter-neural communication, the introspective model is part of the brain's communicative mind myth – the inside place of our memories, imagination, thoughts and feelings. Brain-sign theory explains straightforwardly why 'we' appear to introspect. 'We' are played out by the causal brain, not just as the story, but the activity itself. Language and introspection, of course, are applied to categories-of-interaction. "How do I feel today? I'm sad."

4.5.4. *The science of narrative, and the foundations of science*

Why can it be difficult to change scientific theories? Why is it so difficult to explain some concepts? Paul Churchland states it: "Theories often – indeed, usually – take years of hard work and practice to grasp and internalize, precisely because there is no analog to the physical virus entering the body [or ideas to the mind!], pill-like or bullet-like at a specific time and place. Instead, a vast reconfiguration of the brain's synaptic connections is necessary in order to imprint the relevant conceptual framework on the brain, a reconfiguration that often takes months or years to complete" (2007, p. 5).

When an 'explanation' describing the world is established (a narrative), as mentalism, changing it to a neural science (brain-sign theory) involves reconfiguring a vast network, or series of networks, and other brain conditions. A new causal orientation occurs, with resulting brain-signs signifying the supposed account and understanding of the theory.

The change does not result from a subject (mentally) knowing the failure of current theories, but from the inadequacy of the brain's causal orientations to the world. The brain (without help from consciousness) can generate a new theory with a more effective causal orientation, which is then neurally interpreted as brain-sign and sharable via language. It is no easy process. Geocentrism to Heliocentrism took over a hundred years to be generally 'accepted' – structurally mimicked is more apposite – once Heliocentrism was posited.

The 'sense' of accepting a theory (category-of-interaction) results from the newly established network which, in the case of a specific scientific model, is set in a wider structure of causal orientation to the world, and can therefore be homologous. 'Accepting' a theory is associated with a 'sense of rightness', both of which are decomposable into bio-physical drivers.

For Geocentrism to Heliocentrism, for example, the 'rightness sense' resulted from consistency with the general emergence of the physical sciences, based on systematic brain/world engagement for structures of cause and effect under experimentation. The foundational bio-physical driver for science was survival, as command of the universe's organization and operation. The causal orientation of the brain was (is) interpreted as categories-of-the-world – in particular sun, planets and gravitational force.

Rejection of Heliocentrism, on the other hand, resulted from neurally encountered evidence, i.e. the movement of the sun and stars across the heavens (categories-of-the-world), reinforced by a causal narrative of divine power – networks neutrally interpreted as the prime mover of the universe, the divine nature of the celestial realm, and an association with human ('supposed') purposes. One category-of-interaction as 'felt' was retaining the causal orientation of the group, which is decomposable into subsets of bio-physical drivers, including the security of the group, and avoidance of threat from the group. But, as a means of practical survival, it was less homologous with the emerging categories-of-the-world of scientific development, which accounted for its eventual demise. Proof, however, is always provisional, as brain-sign theory demonstrates.

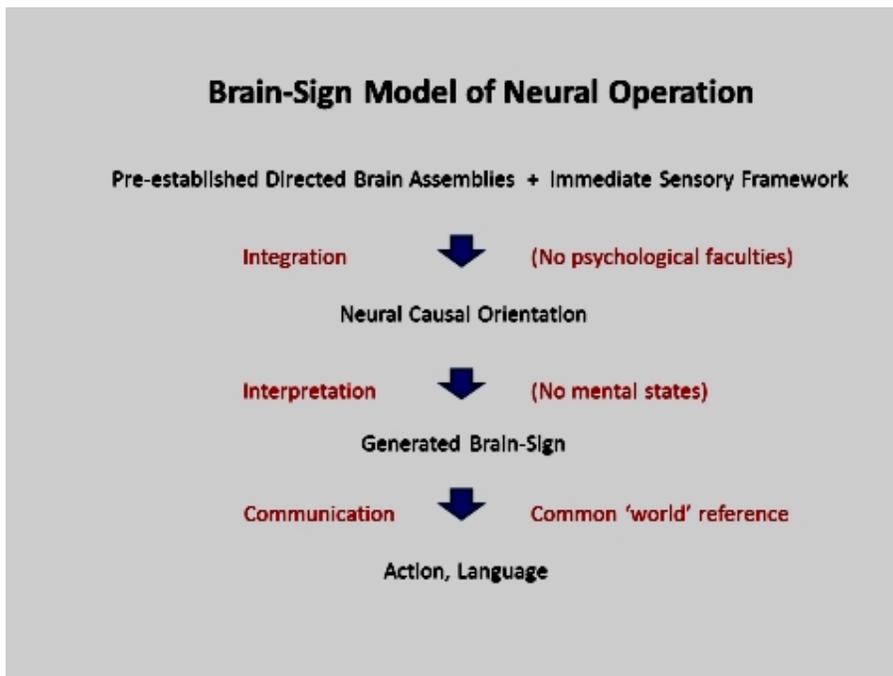


Figure 1. Brain-Sign model for basic neural operations.

5. CONCLUSION

Humans supposed the sun circled the earth, or species occurred by spontaneous generation, because their brains failed to develop models that more correctly predicted the physical behavior of the world, viz. Heliocentrism and Evolution. But as Kuhn (1996) pointed out, it is not that the wrong model may not work somewhat, or that it has an apparent negative impact on daily life.

The functional problem with psychology, as noted at the beginning, is that it is not a science. While the brain is generally regarded as the physical cause of our actions, psychological states, purportedly causal, have not been located in the brain. Cognitive neuroscientists generally assume that psychological states will be locatable, and therefore can be used to explain the causal processes of the organism. After all, they work well enough in daily life. But as this paper has demonstrated, the brain-sign model has a more legitimate claim to be a scientific account of brain operation than psychology.

1) Brain-sign states are not, as in popular usage, perceptions, feelings, thoughts, which are projected onto the brain without scientific explanation.

2) Brain-sign states fit within current biology; they have the physical properties of signs. They will be found to be part of brain fabric.

3) Inter-neural communication is a major topic of neural science. Without communication there would be no organisms as complex as humans. But the topic lacks physical explanation or terminology. Psychological terminologies are certainly a product of brain operation. But psychology cannot form a science of neural communication. Brain-sign theory describes how all theories derive from the brain including itself.

4) Knowledge, and its use, have been a major problem in philosophical and scientific theorizing for millennia. It has seemed that humans depend upon it, for we (supposedly) experience having it. Brain-sign theory demonstrates that there is no knowledge and no need of knowledge. It does so within a framework deriving from the bio-physical operation of the brain and organism. In other words, the phenomena which have been taken as knowledge perform a different role (more complex, more scientifically viable) than the simplistic notion of consciousness. Indeed consciousness, as a theory, obstructs the development of neuroscience as a science.

ACKNOWLEDGEMENT

I must thank Anne Kollar for scrupulous attention and advice on the expression and lucidity of this paper.

REFERENCES

- Anderson, M.L. (2014). *After Phrenology: Neural Reuse and the Interactive Brain*. A Bradford Book, MIT Press.
- The Atlantic* (13 April 2016). <http://www.theatlantic.com/science/archive/2016/04/brain-prosthetic-allows-paralyzed-man-to-move-his-hand-again/478026/>
- Baars, B. (1996). *In the Theater of Consciousness: The Workspace of the Mind*, Oxford University Press.
- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 8, 227-287.
- Chalmers, D. (1995-7). Facing up to the Problem of Consciousness. In *Explaining Consciousness: The Hard Problem*, J. Shear (Ed.). A Bradford Book, MIT Press (and the *Journal of Consciousness Studies*, 9-32).
- Chalmers, D. (1996). *The Conscious Mind*. Oxford University Press.
- Chomsky, N. (1998). *Language and problems of knowledge*. Cambridge, MA: MIT Press
- Churchland, P.M. (2007). *Neurophilosophy at Work*. Cambridge University Press.
- Churchland, P.S. (2013). *Touching a Nerve: Our Brain, Our Selves*. W.W. Norton.
- Clapson, P. (2001). Consciousness: The Organismic Approach. *Neuro-Psychoanalysis*, 3, 203-220.
- Clapson, P. (2014). Knowledge, Science and Death: The Theory of Brain-Sign. *Activitas Nervosa Superior*, 56 (4):105-120. <http://www.activitas.org/index.php/nervosa/article/view/184/205>
- Clapson, P. (2017). The Theory of Brain-Sign: A New Model of Brain Operation. In press in the volume: *The Human Sciences After the Decade of the Brain*, J. Leefmann, E. Hildt (Eds.), Elsevier.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36, 181-204.
- Clark, A. (2016). *Surfing Uncertainty: Prediction, Action and the Embodied Mind*. Oxford University Press.
- Craig, A.D. (2009). How do you feel - now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, 10, 59-70.
- Crease, R. (2008). *The Great Equations: Breakthroughs in Science from Pythagoras to Heisenberg*. W. W. Norton.
- Critchley, H.D. (2004). Neural systems supporting interoceptive awareness. *Nature Neuroscience*, 7, 189-195.
- Dale, R., Fusaroli, R., Duran, N.D., & Richardson, D.C. (2013). The self-organization of human interaction. *The Psychology of Learning and Motivation*, 59, 43-96.

- Damasio, A. (1999). *The Feeling of what Happens*. William Heineman.
- Dehaene, S., Changeux, J.P. (2011). Experimental and Theoretical Approaches to Consciousness. *Neuron*, 70, 200-227.
- Dennett, D.C. (1991). *Consciousness Explained*. Little, Brown & Company.
- Dennett, D.C. (1996). *Kinds of Minds: Towards an Understanding of Consciousness*. Weidenfeld and Nicolson.
- Dennett, D.C. (2007). Philosophy as Naïve Anthropology. In *Neuroscience and Philosophy: Brain, Mind and Language*, pp. 73-96. Columbia University Press.
- Edelman, G. & Tononi, G. (2000). *Consciousness: How Matter Becomes Imagination*. Penguin.
- Fodor, J.A. (1990). *A theory of content and other essays*. Cambridge, MA, MIT Press.
- Fodor, J.A. (2000). *The mind doesn't work that way*. Cambridge, MA, MIT Press.
- Fodor, J.A. (2007). Why pigs don't have wings. *London Review of Books*, 29, 19-22.
- Freud, S. (1969). *An Outline of Psycho-Analysis*. The Hogarth Press.
- Frith, C. (2007). *Making up the Mind: How the Brain Creates our Mental World*. Blackwell.
- Gaukroger, S. (1990). 'Introduction' to Arnaud's *On True and False Ideas*. Manchester University Press.
- Gazzaniga, M. (Ed.) (2009). *The Cognitive Neurosciences*. (4th ed.) MIT Press.
- Gazzaniga, M.S. (2012). *Who's in Charge: Free Will and the Science of the Brain*. Robinson.
- Gazzaniga, M., Ivry, R.B., Mangun, G. R. (2013). *The Cognitive Neurosciences: The Biology of the Mind*. MIT Press.
- Gazzaniga, M.S., Mangun, G.R. (Eds.) (2014). *The Cognitive Neurosciences*, 5th edition. MIT Press.
- Griffiths, P.E. (1997). *What Emotions Really Are*. The University of Chicago press.
- Graziano, M.S.A. (2013). *Consciousness and the Social Brain*. Oxford University Press.
- Heidegger, M. (1962). *Being and Time*. J. Macquarrie and E. Robinson (Trans.). Basil Blackwell.
- Hergenhahn, B.R., Henley, T.B. (2014). *An Introduction to the History of Psychology*, 7th edition, Wadsworth, Cengage Learning.
- Huxley, T. (1874). On The Hypothesis that Animals Are Automata. In *Collected Essays. 1893-1894*. London: Macmillan.
- Jackendoff, R. (1987). *Consciousness and the Computational Mind*. MIT Press.
- James, W. (1894). The physical basis of emotion. *Psychological Review*, 1:516-529.
- Kandel, E (2013). The New Science of Mind and the Future of Knowledge. *Neuron*, 80, 546-560.
- Kashdan, T.B., Steger, M.F. (2004). Approaching Psychological Science with Kuhn's Eyes. *American Psychologist*. May-June.
- Kouider, S., Sackur, J. (2014). Consciousness and its Access Mechanisms. In Gazzaniga and Mangun, 2014, pp. 855-864.
- Kuhn, T. (1996). *The Structure of Scientific Revolutions* (3rd ed.) University of Chicago Press.
- Lange, C. (1885). *Om Sindsbevaegelser*. Kjöbenhavn.
- Lashley, K. (1951). The problem of serial order in behavior. *Cerebral mechanisms in behavior*, L.A. Jeffress (Ed.). Wiley, pp. 112-146.
- Libet, B., Gleason, C.A., Wright, E.W., Pearl, D.K. (1983). Time of conscious intention to act in relation to onset of neural activity (readiness potential). The unconscious initiation of a freely voluntary act. *Brain*, 106, 623-642.
- Louwerse, M. M., Benesh, N., Watanabe, S., Zhanb, B, Jeuniaux, P., Vargheese, D. (2009). The multimodal nature of embodied conversational agents. In N. A. Taatgen, H. van Rijn (Eds.). *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 1459-1463). Austin Texas: Cognitive Science Society.
- Lieberman, P. (2006). *Toward an Evolutionary Biology of Language*. Belknap, Harvard.
- Macphail, E.M. (1998). *The Evolution of Consciousness*. Oxford University Press.
- Merleau-Ponty, M. (1962). *Phenomenology of Perception*. C. Smith (Trans.). Routledge.
- Merleau-Ponty, M. (1968). *The Visible and the Invisible*, A. Lingis (Trans.). Northwestern University Press.
- Metzinger, T. (2003). *Being No One. The Self-Model Theory of Subjectivity*. MIT Press, Cambridge, MA.
- Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, 83:435-50
- Norretranders, T. (1998). *The User Illusion*. J. Sydenham (Trans.). Viking Penguin.
- Oatley, K., Keltner, D., Jenkins, J.M. (2006). *Understanding Emotions*. Malden, MA: Blackwell.
- O'Shea, M. (2005). *The Brain: A Very Short Introduction*. Oxford University Press.
- Pessoa, L. (2013). *The Cognitive-Emotional Brain*. MIT Press.
- Pinker, S. (1997). *How the Mind Works*. W. W. Norton.
- Pinker, S. (2007). The Brain: The Mystery of Consciousness. *Time*, January 29.
<http://content.time.com/time/magazine/article/0,9171,1580394,00.html>

- Rizzolatti, G., Fadiga, L., Gallese, V., Fogassi, L. (1996). Premotor cortex and recognition of motor actions, *Cognitive Brain Research*, 3, 131-141.
- Schopenhauer, A. (1985). *On The Freedom of the Will*. K. Kolenda (Trans.). Basil Blackwell.
- Seager, W. (1999). *Theories of Consciousness*. Routledge.
- Searle, J. (1992). *The Rediscovery of the Mind*. A Bradford Book, MIT Press.
- Searle, J. (1998). *The Mystery of Consciousness*. Granta Books.
- Searle, J. (2014). Interview. *The New Philosopher*. Issue no. 2.
<http://www.newphilosopher.com/articles/john-searle-it-upsets-me-when-i-read-the-nonsense>
- Seth, A.K. (2013). Interoceptive inference, emotion and the embodied self. *Trends in Cognitive Sciences*. 17, 565-573. <http://www.cell.com/trends/cognitive-sciences/fulltext/S1364-6613%2813%2900211-8>
- Smith, R. (2013). *Between Mind and Nature*. Reaktion Books Ltd.
- Swaab, D.F. (2014). *We Are Our Brains: From the Womb to Alzheimer's*. Allen Lane, The Penguin Press.
- Tomasello, M. (2014). The Ultra-Social Animal. *European Journal of Social Psychology*. 44, 187-194.
<http://onlinelibrary.wiley.com/doi/10.1002/ejsp.2015/epdf>
- Tononi, G., Koch, C. (2015). Consciousness: here, there and everywhere? *Royal Society, Philosophical Transactions B*. <http://rstb.royalsocietypublishing.org/content/370/1668/20140167/>
- Ward, J. (2012). *The Student's Guide to Social Neuroscience*. Psychology Press.
- Wegner, D.M. (2002). *The Illusion of Conscious Will*. MIT Press.