



## Durham E-Theses

---

### *The world without knowledge: The Theory of Brain-Sign*

CLAPSON, PHILIP,JOHN

#### How to cite:

---

CLAPSON, PHILIP,JOHN (2012) *The world without knowledge: The Theory of Brain-Sign*, Durham theses, Durham University. Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/3560/>

#### Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

# **The world without knowledge: The Theory of Brain-Sign**

Philip Clapson

## **Abstract**

This thesis proposes the theory of brain-sign should replace the theory of consciousness, and mind generally. Consciousness is first described, and it is evident there is neither a consistent account of it, nor an adequate rationale for its existence beyond its (tautological) self-endorsing characterisation as knowledge. This is the target of the critique. Human history and human self-appraisal have assumed knowledge as an (ontologically) intrinsic facet of human make-up, and human disciplines (of ‘thought’) have relied upon it. On the modern supposition that the universe is purely physical, it is widely assumed that science eventually will demonstrate that consciousness, and thus knowledge, are physical. But it is shown here (theoretically and empirically) that as theories of the physical universe, consciousness and thus knowledge are prescientific, incoherent and unmotivated.

Nevertheless, there is a brain phenomenon requiring explanation. That explanation is given here by a new account: *the theory of brain-sign*. Brain-sign is developed as a scientific theory, i.e. directly from the physical universe, and with a fundamental biological role as the communication mechanism for (opposite) organisms in dynamic (i.e. uncertain or imprecise) collective operation. The grounding principle is that the brain is an action organ, not a knowledge organ. The detail and structure of brain-sign is expounded, and its success is contrasted with the inadequacy of consciousness. Humans, and other organisms, are placed in the physical world, and in the context of evolutionary theory. Thus brain-sign theory founds an approach to brain science, but with the result that human self-conception alters fundamentally, as does the relation between the organism and its disciplines, including science itself. Brain-sign theory shows why there ever was a theory of consciousness, or indeed a theory of anything. However, from the nature of our being in the world, it also demonstrates it cannot claim to be true.



**The world without knowledge:  
The Theory of Brain-Sign**

**Philip John Clapson**

*PhD Thesis*

**Department of Philosophy, Durham University**

**2011**



# **The Theory of Brain-Sign**

## **Contents**

<b>Acknowledgements</b>	11
<b>Preface</b>	13
<b>1.0 Introduction</b>	17
<b>2.0 What is consciousness?</b>	28
2.1 The fundamental premise: knowledge	29
2.2 The I, self or ego	31
2.3 Interlude	33
2.4 Inner and Outer	33
2.5 Conscious experience	34
2.6 Illumination	40
2.7 Consciousness, the unconscious & non-conscious	42
2.8 Intentionality	44
2.9 Memory	46
2.10 Imagination	48
2.11 Feeling, sensation and mood	50
2.12 Free Will	51
2.13 Summary	52
<b>3.0 What is wrong with consciousness?</b>	54
3.1 Representation and correspondence	54
3.1.1 Representation from Descartes to the Empiricists	55
3.1.2 Locke's absence of justification	59
3.1.3 The problem of covariance and misrepresentation	60
3.1.4 Slezak's critique	61
3.1.5 Slezak's unintended destruction of consciousness	65
3.2 Illumination	69
3.2.1 The illuminating topic of colour	69

3.2.1.1	The evolutionary factor	71
3.2.1.2	Rey, Ross, and the unacknowledged end of mentalism	74
3.2.2	Summary	77
3.3	Empirical findings	77
3.3.1	The consciousness structure	78
3.3.2	The breakdown of the brain phenomenon	81
3.3.3	Milner and Goodale: Two visual streams of the brain	82
3.3.4	Language	87
3.3.4.1	Language and neuroscience	88
3.3.4.2	Language and action	91
3.3.5	Libet	95
3.4	Summary	96
<b>4.0</b>	<b>The predicament, &amp; the mentalist error of knowledge</b>	97
4.1	Prelude to the predicament	97
4.1.1	The predicament exposed	99
4.1.2	The predicament considered	100
4.1.3	The predicament and free will	104
4.2	Bennett & Hacker, and the ontological issue in neuroscience	105
4.3	Paul Churchland's connectionist knowledge	110
4.4	Conclusion	116
<b>5.0</b>	<b>Introducing the theory of brain-sign</b>	118
5.1	Dualism lies not in ourselves but in our theories	118
5.2	A brief encounter with history	120
5.2.1	Pragmatism	121
5.2.2	Phenomenology	125
5.2.3	Summary	133
<b>5.3</b>	<b>The scientific model of brain-sign: What biological problem is to be solved in the brain? The super-organism</b>	135
5.3.1	Preliminary orientation – the natural and scientific dispositions	135
5.3.2	The grounds of the biological model	137

5.3.2.1	The literature confusion	137
5.3.2.2	Communication supersedes the individual	139
5.3.2.3	Communicative modalities – the biological differentiation	142
5.3.2.4	Signification – static and dynamic communication	142
5.3.2.5	Signification of affective states and reference in the literature	143
5.3.2.6	Static communication	143
5.3.2.7	Dynamic communication	146
5.4	Advantages of the brain-sign account	148
5.4.1	The explanation of explanation	151
5.4.2	Direct reduction to the physical	151
5.4.3	Separation of scientific questions	152
5.4.4	Evolutionary continuity	152
5.4.5	Dissolution of the inner and outer conundrum	153
5.5	The integral nature of brain-sign	153
<b>6.0</b>	<b>Characterising brain-sign</b>	<b>155</b>
6.1	The natural and scientific dispositions	155
6.2	Brain-sign and the virtual array	156
6.3	The instantaneous nature of the physical universe	158
6.4	Our dual nature: Causality and communication	160
6.5	Removing the need for causal conscious representation	164
6.5.1	The connection between brain models	167
6.6	Introduction to two fundamental elements of brain-sign: categories-of-the world & categories-of-interaction	167
6.7	The content of brain-sign	171
6.8	Control and duration	173
6.9	Subject, self, ego	176
6.10	Memory and imagination reinterpreted as physical states	179
6.11	The scientific explanatory power of brain-sign theory vs. mentalism	184
6.12	Development of categories-of-the-world and -interaction	187
6.12.1	The ontological contrast of brain-sign categories vs. mental states	187

6.12.2	The general contrast of categories-of-the-world and -interaction	188
6.12.3	What mental state terms are as brain-sign categories	191
6.12.4	Simulation theory and theory theory	192
6.12.5	Brain-sign eliminates the subjective/objective distinction: Consider pain	194
6.12.6	Brain-sign terminology will not replace mental terminology generally	195
6.12.7	The question of terminology	195
6.13	Music	197
6.14	What do deficit and illusion show?	202
6.14.1	Deficit	202
6.14.2	Illusion	203
6.15	Conclusion	205
<b>7.0</b>	<b>Language</b>	207
7.1	Preliminary considerations	207
7.1.1	The general situation	207
7.1.2	The relevance of Saussure	208
7.2	The literature on language	210
7.2.1	Thinking and communication	210
7.2.2	Chomsky's approach to thought	214
7.2.3	Thought <i>via</i> meaning and understanding	216
7.3	The social factor & learning language	218
7.4	Lieberman's attack on Chomskian language	221
7.5	Language as action	224
7.5.1	How language functions	227
7.5.2	Communication with animals	226
7.5.3	Intelligence lies in action	226
7.6	The operation of language	227
7.6.1	Neural collectivity	228
7.7	The nature of language, and our fundamental limits as biology	229
7.8	A recreational coda	232

7.8.1 A review of the two forms of joke explanation	235
<b>8.0 The organism without consciousness. Science and other human disciplines.</b>	237
8.1 Theories (leading to a consideration of science)	237
8.1.1 First stage: Current misrepresentation of the nature of the problem	240
8.1.2 Second stage: A current status in the Philosophy of Science	242
8.1.3 Third stage: Uniting in review stages one and two	245
8.1.4 The resolution of the science question	247
8.1.5 Error	249
8.1.6 Adjunct topics	250
8.1.6.1 Realism/Instrumentalism	252
8.1.6.2 Thomas Kuhn	252
8.2 Afterword	255
<b>9.0 Considering ourselves</b>	256
9.1 Light	258
9.2 Our reality (by contrast with mentalism)	258
9.3 Transparency; opacity	261
9.4 Heaviness; lightness	263
9.5 The absence of memory	266
9.6 What reflection really is	266
9.7 The brain is an alien	268
<b>10.0 Conclusion</b>	272
10.1 Summary	272
10.2 Future work	274
<b>Bibliography</b>	280

***The copyright of this thesis rests with the author. No quotation from it should be published without the prior written consent and information derived from it should be acknowledged.***

## **Acknowledgements**

I considered the problems of mind, and specifically consciousness, over many years. Eventually I found a way to address the problems with it. My supervisor at Durham, Jonathan Lowe, generously allowed me to develop my solution at length, and then reduce and refine it to the PhD thesis. I would like to thank Jonathan for many interesting discussions on the topic.

I would like to thank Sophie Gibb for initial discussions on the subject matter. I would also like to thank Wolfram Hinzen for extensive comments on the early chapters of the thesis.

The thesis was completed with Nick Zangwill as supervisor. I wish to thank Nick for comments on the entire text, and in particular for those which have improved the expression and clarity of the topic as a whole.

Attending Durham University allowed me to write out in detail the theory of brain-sign. I am very grateful this opportunity was afforded to me by the Department of Philosophy.



# **The Theory of Brain-Sign**

## **Preface**

The thesis is in two parts. The first is a critique of the notion of consciousness; the second presents a new (wholly physical) scientific theory of the brain phenomenon misconceived as consciousness, *viz.* The Theory of Brain-Sign. The critique of consciousness must come first so that a need for an alternative theory becomes not merely necessary but pressing. But this involves a delay for the reader in acquaintance with the theory. This Preface, therefore, introduces the theory very briefly. The full extent, detail of derivation, and justification must be left to the subsequent text.

A number of difficulties attend the topic, two of which are that the theory is elaborate, and it undermines our natural disposition towards ourselves. Nonetheless it offers a resolution of ancient and seemingly intractable issues.

Although the mind-body problem is not new, over the last sixty years it has become a major concern not only for philosophy but neuroscience as well. There is no universally accepted scientific solution. The theory of brain-sign takes the position that the mind-body problem – how the mind can be part of a physical universe which conforms to physical laws – is misconceived. This is because the mind does not exist. Therefore the problem disappears because what replaces the mind is not only unproblematically physical, but fulfils a genuine biological function. Brain-sign is the name I give to that replacement condition and function.

The fundamental proposition is that the brain is a causal organ, not a knowledge organ. The theory therefore assumes the brain does not know anything: it is a purely physical object. However, brain-sign, in every apposite organism's brain, is the signification of the cause of the brain's current causal orientation towards the world. Brain-sign derives from the brain's causal orientation toward the world by a neural mechanism as yet to be determined. However, it is the causal orientation that provides the foundation for the brain phenomenon, as a sign, to exist at all.

The function of brain-sign is to signify the brain's causal orientation for the brains of other organisms in what is termed dynamic (imprecise or uncertain) collective activity. A basic example (as introduced in the text) is two lions hunting a gazelle. The common object is the gazelle, and each brain will signify the gazelle 'image' as that common target because, in

the hunt process, the gazelle 'image' in each brain derives from the causal orientation (as purely physical states) of each lion's brain, i.e. the result in the brain of the causal interaction with the gazelle. The gazelle *is* the cause of the brain's causal orientation. Thus the hunt is effected by both the causal orientation of each brain, and their mutual signification of their target, the 'image' of the gazelle. Sustained loss of the gazelle 'image' means the hunt has ceased to happen.

In a similar way in looking at a tree, you and I (our purely physical brains) can mutually react to it because both our brains are causally orientated towards it, the 'image' of the tree being the common sign for our (brains') collective activity. The 'image', the common sign, just is physical neural communication.

The contrast of consciousness and brain-sign is thus established. While the conscious image appears to arise without any non-tautological biological rationale (we are conscious therefore we know: we know because we are conscious), brain-sign arises directly from the brain's causal orientation, and facilitates collective action. With brain-sign there is no knowledge of what the image is about that causes action, thought or feelings. In other words, the 'image' does not function for us, the individual organism. Its biological role is entirely exhausted in its communicative function.

The difficulty in establishing consciousness in the brain has two components. The first is: Where and how does it exist in the brain? The second is: What is consciousness, and therefore knowledge? Since there is no agreed answer to the second question (nor scientific account of it), the pursuit of consciousness in the brain is problematic at its origin. For brain-sign, on the other hand, the only question is: Where and how does it exist in the brain? This is because signs are a biological commonplace.

It is proposed, therefore, that brain-sign theory is a new fundamental scientific theory because the nature of its signification does not involve a conscious (or unconscious) recipient. (No lion's brain could look into another's to 'see' the sign.) Therefore there is no question of its function being (in mentalist language) intentional. The 'image' signifies solely as a necessary part of the physical/causal process of collective action. Nor does the 'image' involve qualia. (The function is perhaps hinted at neurally by the discovery of mirror neurons, although these do not constitute a sign.)

Straightforward signification is not eliminated from the universe. But one organism's external signs to another do not require consciousness in the recipient. They only require that the causal/physical orientation of the recipient organism is altered by the signs. Thus use of the word 'communication', which applies both to external signs and the internal signing of brain-sign, remains appropriate, since in today's usage, communication can apply both to that involving knowledge (requiring consciousness), and the mere alteration of physical states. In fact, brain-sign theory eliminates the first kind of communication because it eliminates consciousness and thus knowledge. The same account applies to use of the word 'information'. It can be used evoking the notion of knowledge, but it is also used purely for the altering of physical conditions.

However, the elimination effected by brain-sign theory is not to be confused with other eliminativisms. Brain-sign theory redefines the condition and function of the brain phenomenon previously understood as consciousness. The result is that we must 'understand' ourselves in an entirely different way – although there is no such thing as understanding in fact, for it too has been eliminated.

Why do we feel we do know? Indeed, why is the sense that we know the justification for consciousness? The answer is straightforward. It is because all components of brain-sign (including and beyond 'seeing') perform the biological function of communication with others in collective action. In action, two or more people are one physical unit. We cannot see this is the case, not because we cannot see all the physicality involved, but because we do not see at all. Thus we 'sense' we see because the conviction of seeing is part of the communicative structure. Hardly could our brains communicate by what we take as seeing if we were unconvinced that we do see. But our conviction is not an illusion. It is (as brain-sign) our biological function. It is the *theory* of consciousness, that we do see, which is an illusion, or better expressed, a mistake.

It will be evident that with the prime function of consciousness eliminated, *viz.* knowledge, we know nothing. This has two convenient consequences. Firstly, since consciousness and thence knowledge are unknowns, we do not have to legitimise them before we have the foundations of a brain science. Secondly, the condition of brain-sign, being physical and deriving from the causal orientation of the brain, means that all biology is cogently identified as belonging to the physical universe. However *de facto* we do not know this. Thus the boundaries of our existence in the physical universe are established – according

to the theory. Epistemological difficulties do not result from the mind being in the head while the world is out there: epistemology itself is eliminated. Communication for organisms/brains without mental states is, however, explained.

It remains only to establish the context of the thesis itself. Brain-sign theory, as here presented, must be seen as the first step in an extensive and staged effort of investigation. Its prime motive is as a propaedeutic for neuroscience. It seeks to lead neuroscience away from the search for consciousness in the brain, and to point to more fruitful lines of investigation. It does this here only by changing our perspective on what the brain phenomenon is structurally and functionally. There can be no attempt to characterise the contents of the brain phenomenon physically in the brain because, for this to happen, brain science would already have to be developed beyond its current state. It is hoped that the development of brain science will involve brain-sign theory because what it does say currently is adequately convincing.

One aspect of this is that mentalist terms, as for example seeing, are generally retained during the explication, but with inverted commas around them: ‘seeing’. This has the positive aspect that the reader knows what is being pointed at. The disadvantage is that what is being overcome is, in a sense, retained. This is an unfortunate consequence of this early stage of the project. But as said at §6.12.7: ‘When the *solar* system was identified, planets retained their geocentric names, though they were no longer the entities previously supposed, nor conformed to the geocentric model in behaviour. Thus there is a precedent for retaining the names provisionally.’

More fundamentally, however, the function of language itself is redefined in brain-sign theory. Words do not mean anything, but between organisms their physical conditions (compression waves and electromagnetic radiation) function to alter the causal orientation of the recipients’ brains. The target for scientific investigation of language does not lie in trying to explain meaning, but in the analysis of how causal neural orientation is altered by the physical conditions. Words ‘read’ or ‘heard’ are themselves brain-signs and causally inert for the individual organism: they function as all other components. Therefore any use of words must be contextualised by brain-sign’s approach to language itself. Again, the development of this lies in the future development of brain science. Ultimately, as with all scientific theories, the theory’s validity will be tested by empirical methods.

# The Theory of Brain-Sign

## 1.0 Introduction

The last four hundred years have shown a transformation of human existence by the rise of science, and the scientific approach. Yet science has found one topic intractable. The nature of mind, and specifically consciousness. Since this is so significant for us, it is striking so little progress has been made. On the one hand we appear to know about the physical mechanics of being born and dying, as the medical and biological sciences describe them. On the other we have no equivalent knowledge as to why we do what we do, feel what we feel, and what the point of these to our existence might be. This is because the terms of such knowledge, viz. the ontology of our mental life, is not merely in a state of controversy, but apparent impenetrability. Despite the assurance of many writers on the topic, controversy is the appropriate word, for there are no generally accepted principles. A few examples will illustrate this.

‘Neuropsychology is in a conceptual morass. Neuropsychologists seek to study the relation between brain and mind, but without really addressing the status of these two constructs, or what potential form the relation between them might take.’

These are J. Graham Beaumont’s words in a *Dictionary of Neuropsychology* which he has edited (Beaumont *et al.*: 1999; 527). It might seem reasonable to ask, in the light of this judgement of the editor, what information exists in the 816 pages of the dictionary. It would hardly seem there is a science to be discussed when the status of the two essential constructs, brain (neuro-) and mind (-psychology), are not addressed, let alone what relation exists between them. If mechanics did not consider what status force, acceleration or friction have, there would be no mechanical science.

Here is a statement from a widely used student text on neuroscience.

‘Exactly how the parallel streams of sensory data are melded into perception, images and ideas remains the Holy Grail of neuroscience’ (Bear *et al.*: 2006; 421).

The content of brain and mind are specified for the student (sensory data and perception): yet what they are, and what goes on between them, is regarded as the Holy Grail. It seems implausible that neuroscience could claim to be a science when its foundational stuffs are

remote as the Holy Grail. Yet routine work on patients' brains is carried out in hospitals across the planet. How is this to be reconciled with the remoteness of the Holy Grail?

Psychologists regard the experience of their experimental subjects as a valid category. Chris Frith, looking at fMRI scanning images, for example, states that: 'By measuring brain activity, we have an objective confirmation of...mental events' (2007; 13). But if Beaumont is correct, then Frith's statement appears incorrect. For Frith can only be *supposing* there are mental events if the status of mental events as a construct is not addressed.

Indeed Frith, in neuroscience mode, goes on to say: 'I could probably find a neuron in [the] brain that only responds when...experienc[ing] the color blue. But...the [brain] activity is not blue' (15). He identifies a fundamental difference between what is indicated in an illuminated fMRI scan of the brain, and what he refers to as a mental event (experiencing the colour blue). He states it: 'What the brain imaging experiments reveal so starkly is the seemingly unbridgeable gap between objective physical matter and subjective mental experience' (*ibid.*).

But Beaumont's point is more precise. It is that the *construct* brain and the *construct* mind have not been addressed. Frith considers himself a scientist and positions himself accordingly. 'I am not a philosopher. I do not expect to persuade people of the truth by the power of argument. The only arguments I accept come from practical experiments' (*ibid.*). But this is at variance with his own account, for his experiments reveal an 'unbridgeable gap'. That kind of gap is what interests philosophers.

However it would not be correct to suggest that philosophers are concerned with the issue Beaumont has raised. Most philosophers do not address the status of the *constructs* brain and mind. They take for granted that brain and mind exist. Their concern is with how the two can be reconciled as such. Beaumont, on the other hand, has questioned whether the constructs *themselves* are correct. This question precedes the arguments of most philosophers.<sup>1</sup>

Over the last fifty years the literature on the topic of brain and mind has greatly expanded. But Beaumont's few words have raised a fundamental issue. Is the literature

---

<sup>1</sup> This does not mean the mind is considered unproblematic. E.g. Peter Slezak: 'The recurrence of certain deep perplexities about the mind is a systematic and pervasive pattern' (2001; 940). 'The suspicion of deep conceptual problems at the heart of philosophy and psychology is...justified' (944). We shall come to these problems.

adequately addressing its subject matter? Is the topic discussed in the right way? On the surface the issue seems easy to point to. The brain appears to be a mass of physical stuff: neurons, neurotransmitters, peptides, etc. The mind, consciousness specifically, is (e.g.) the experience of redness in the sky. How is it possible the experience of redness is the same as the state of a, or many, neuron(s)?

This kind of example is commonly referred to when posing the problem of brain and mind. But Beaumont states that this way of specifying the issue could be wrong. He conjectures the fundamental constructs, so expressed, may lead us astray.

The natural reaction is scepticism. We are perfectly acquainted with the world of physicality for it appears before us when we open our eyes. At the same time we are immediately acquainted with our own mental states, the seeing of the world out there and our pleasure at the redness in the sky. It seems impossible to suppose, given the world *is* right there in front of us, that the world is *not* there; or, on the other hand, that this feeling of pleasure at the redness of the sky does *not* occur as we feel it. To question these would be to question our very existence. Thus (we suppose) the brain *must* be as we see it: neurons, neurotransmitters, peptides; and the mind, including this pleasurable experience of the sky, must also be as we take it. Even so, according to Beaumont, there must be something wrong with our certainty, for the fundamental constructs are not addressed.

How would we address them? What does *addressing* them mean?

Let us put this another way. What would we be doing by addressing the constructs? Philosophy or science?

At the dawn of the scientific age, when Copernicus, Kepler and Galileo addressed the character of the universe by questioning the Aristotelian/Ptolemaic construct, what were they doing? The answer must be philosophy *and* science (in modern taxonomy). Not only did they identify how the universe behaved lawfully in a scientific/functional sense, they were concerned with the ontology – the nature of being – of the relevant entities. They established the centrality of the sun in the solar system (helio- not geo-centrism), and with Galileo's observations, gone were the heavenly spheres and transcendently perfect celestial points of light. Planets and stars were gross material entities.

But there is another sense in which philosophy and science conjoined. What had to be overcome *for* science, was the observer's position. Looking upwards and outwards was not

adequate for discovering either how the universe operated or what its entities were. Our 'natural' tendency to assume what is from how things appear to us inhibited scientific and ontological progress. In the post-medieval/post-renaissance world, the notion of heliocentricity took well over a century for general acceptance both because of religious denial, and straightforward inconceivability. It was not until Newton's specification of gravity and laws of motion, that motion in the heavens conformed to a unifying theory with motion on earth. The whole universe became one kind of thing.

The confusion about the earth and heavens, lasting roughly two thousand years, ended. The resulting shock was great. Where, now, was God? What was the point of life if man's unique significance was gone? Could life be no more than 'a walking shadow'?

A direct link between cosmology and mind was Descartes. For it was Descartes' mechanical view of nature that clearly conceived the inertial condition: particles at rest unless external force applied, which Newton inherited. This same Descartes conceived the *cogito*, founding the mind upon the I that thinks, and knows it thinks. It was a concept of personal independence. The rational mind was a bold assertive antidote to life as 'a walking shadow' in the loss of old certainties. However, the *cogito* was a problem from the start precisely because it could not be reconciled with the brain, which was part of the mechanical universe. How could the thinking mind act upon physicality if it was not part of that physicality? as Princess Elisabeth of Bohemia asked in her letter.

Going on some centuries, Beaumont raises a more probing question in the ontological spirit of the new cosmology. The new cosmology rejected our (geocentric) observational point of view. Yet Descartes at the same time (historically) established the centrality of our mental point of view in his notion of the mind: '*I think, therefore I am*' (echoing God's self-description: 'I AM THAT I AM'). Was this the crucial mistake? I.e. not so much Descartes' dualism, but the construct of the mind model itself, with a subject-based foundation.

There are two existence claims in Descartes' model. The first goes by almost without acknowledgement. *I exist*. This means there is a subject; it does exist; and I am it. The second is that the means by which the I exists is thinking. Therefore, *de facto*, thinking exists. The subject exists as a thinking thing.

The basis of Descartes' model is, therefore, an existence claim about the mind and its structure, it being self-revelatory. Though questions were raised at its inception by acute

observers (e.g. Princess Elisabeth), the model declares its foundations as unassailable. This approach is extant today. Ned Block says: ‘The starting point for work on consciousness is introspection and we would be foolish to ignore it’ (2001; 203). Block founds his statement upon the assumption that introspective self-revelation exists. With or by the mind we can look internally at our own condition, as well as externally. But may it not be that this assertive stance is the very reason the mind, specifically consciousness, fails to find resolution as a scientific theory? I.e. *can* we look internally? (Or externally for that matter?) To doubt Block’s position would certainly be to question the *construct* mind.

Such questioning can receive an apocalyptic response. Here is a well-known passage from Jerry Fodor.

‘If commonsense intentional psychology really were to collapse, that would be, beyond comparison, the greatest intellectual catastrophe in the history of our species’ (1987; xii).

Commonsense intentional psychology (i.e. explaining our actions by reference to our mental states, e.g. our beliefs and desires) depends upon the existence of consciousness. Psychology presupposes a *psyche*, i.e. consciousness. We are warned not merely against foolishness, but the collapse of intellectual life *per se* if Descartes’ model (at least in foundational terms) is incorrect and there is no *psyche*.

Even so, we are faced with a fundamental anomaly in the literature at large. This concerns the word *knowledge*. Knowledge, of course, depends upon consciousness: con-(with) -scious (science, i.e. knowledge).

In the student text quoted above, is this passage.

‘Perception of any handled object involves the seamless coordination of all the facets of somatic sensory information. For example, the bird in the hand is rounded, warm, soft, and light in weight, its heartbeat flutters against your finger tips, its claws scratch, and its textured wings brush against your palm. Somehow your brain knows it’s a bird even without looking or listening and would never mistake it for a toad’ (Bear *et al.*: 2006; 421).

It is evident why the writers regard the finding of perception, images and ideas in the brain as the Holy Grail. It lies in the word ‘somehow’: ‘Somehow the brain knows it’s a

bird.’ If the brain is purely physical matter, that ‘somehow’ flags the problem of relating the brain as purely physical matter to the knowledge of anything.<sup>2</sup> For *knowledge* is a property of Block’s and Fodor’s *psyche*. Block and Fodor ‘get’ this *psyche* from Descartes, and human cultural history disappearing into unrecorded time. But how the *psyche* (if it exists) could relate to the brain is precisely the problem.

For nearly two thousand years, seekers after knowledge puzzled over the nature, structure and workings of the universe. They had to overturn much to get an elegant and successful solution. They had to find an appropriate model. It would appear, on the evidence so far offered, that one is needed now for neuroscience.

\*

This thesis proposes such a model. It is termed: *The theory of brain-sign*. As the name implies it is a theory of brain states (those misconceived as consciousness), and it proposes these states act for the brain/organism as a sign.

Brain-sign theory concerns what we are, and is not merely something we can know, or learn conceptually (because brain-sign states are *not* consciousness). In understanding brain-sign theory, therefore, one re-interprets oneself in a fundamental way. However, brain-sign theory is conceived as a scientific theory (to be empirically justified).

Science concerns itself with the physical universe, with matter. There is inorganic matter; organic matter; and then a more vague domain which is termed mind or consciousness. The sciences of inorganic matter are many and diverse, from quantum physics to relativity theory, including theories of the origins of the universe. Organic sciences, here specifically forms of biology, are also diverse, and refer back to inorganic matter as the foundation of the behaviour of organisms, or their parts, as biological. Of course these two explanatory domains sustain controversies and different approaches. But the foundations are not under suspicion in the same way as mind or consciousness.

Brain-sign theory proposes there is no emergent (or otherwise existent) domain as mind or consciousness. As a theory of science it is an account concerning the ontology and behaviour of matter. Brain-sign theory does not deny something happens in the brain which

---

<sup>2</sup> Whether the brain knows anything features in Maxwell Bennett and Peter Hacker (2003), and the resulting discussion of those authors with Daniel Dennett and John Searle (Bennett *et al.*: 2007). Specific reference to these texts and positions will be made later.

has been mistaken for consciousness. Indeed, what happens has a vital biological role and functional consequence.

Thus what has to be demonstrated, in a negative sense, is twofold. We must show why the theory of consciousness is wrong, why it is prescientific. This involves scrutinising the various aspects of consciousness theory, including aspects that are seldom challenged, if at all. But also there is the need to show why there *is* a theory of consciousness, and this will turn out to be a particular case of why there are theories of anything.

The project of the last fifty years, to show how consciousness is, is not, or in what way aspects of it could be physical, will be shown to be undermined at the outset because the assumption that consciousness *exists* is wrong. Though there are physicalist accounts of consciousness that attack covert or declared dualist positions as those of Ned Block and David Chalmers, notably those of Daniel Dennett and Paul and Patricia Churchland, brain-sign theory does not align with them because they still assume consciousness does exist. Indeed, it might seem remarkable (and will be discussed at greater length) that, whereas dualists of varying denominations attempt to hold onto the notion of consciousness as a tradition, as it were, those orientated toward science (as a realm of enquiry concerning matter) should not see that consciousness is a dualist concept *simpliciter*. In other words, there is no point in setting out to reduce or make identical conscious states with matter when they belong to the wrong model.

Although philosophers have been custodians of consciousness, their inability to view the concept dispassionately leaves psychologists and neuroscientists, who follow in their train, without guidance. Frith (psychologist/neuroscientist) has not seen that the problem with his statements about what fMRI scans show actually results from his entrenchment in philosophical history. While he is dismissive of philosophers and their (argumentative) endeavours, if there is no consciousness, what is the status of his science?

To arrive at brain-sign theory we must step back from cultural suppositions to physical biology. Consciousness, as a theory of our being, fills in gaps which must be exposed. There is an analogy between the explanatory role of consciousness and the idea of God. It was convenient to suppose the universe was explainable because God created it, and he cared about us, being a universal yet particular (to us) father.

The illusion in the theory of consciousness (we shall find) is even more fundamental: that the world is there to know and feel something about. As has often been pointed out, and is crucial to certain philosophical positions, knowing and feeling are not open to measurement. So as knowing sentient human beings, we *escape* the rigours of science (as we escape our physical limitations in the notion of God).<sup>3</sup> Science is the objective bogy to our unique individuality; and this individuality is what we value most.

\*

*But* (and this illustrates the problem with the topic), the way this has been stated is already somewhat misleading. From a biological stance, we cannot say the idea of God was comforting because both the idea of God and comfort are dualisms. To hold that the idea of God gives us comfort is to suppose there is some kind of causal role an *idea* can perform. Causality requires there is some physical state to be causal. But is an idea physical? Moreover, this causal role operates to enable comfort. Is comfort physical?

But before asking whether an idea can cause anything, we need to ask (as Beaumont has suggested): What is an idea? For to say an idea would have to be a state of the brain is no more illuminating than Frith's experienced colour blue, or our seeing a red sky. Even before we consider the brain, we lack any clear idea of what an idea is. But in so saying, the nature of the problem is immediately evident. By ignoring our ignorance of what an idea is, we presuppose there must be a self-evident explanation of an idea, because ideas are what we use to explain things. But how could we use the very thing we are trying to explain (an idea) for an explanation of what we are trying to explain (ideas)? Explanation is what we do with ideas, but how *it* works requires some other (more ontologically fundamental) terms in which *it* can be explained.

In a similar way, it would be misleading to say that consciousness gives us the illusion that we know and feel. The notion of *illusion* stems from the idea that only the brain can be causal, because it is what is physical.<sup>4</sup> Therefore to suppose our conscious (or mental) states (e.g. our knowing or feeling) are also causal is to suppose there are two sources of causality,

---

<sup>3</sup> Even a writer on science such as Robert Crease can say: '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; 64).

<sup>4</sup> The 'illusion' position, in which consciousness is both assumed and its powers denied, can be seen in many places, particularly Dennett, to whose work we will return. But e.g. Norretranders (1991), Claxton (1999), Wegner (2002).

our physical brains and our mental states (e.g. knowing and feeling) – and this cannot be, and is termed causal over-determination. Therefore, the argument goes, our conscious states must be illusory in terms of causality, and some theorists go on to assert that conscious states are epiphenomenal.

But an illusion can only occur *to* someone. Thus to suppose the causality of consciousness is an illusion is already to have posited the conscious person *to whom* the illusion occurs. Thus, in a similar way to explanation itself being incapable of explaining explanation, it is not possible to suppose that consciousness is causally illusory because in *knowing* it is illusory it is already taken as causal, otherwise knowing is a contradiction of itself. I.e. if I know something, then I can act knowing it; but if I know my knowing is causally illusory, then I (as conscious) have become irrelevant to the universe that I know, including myself. This appears to involve a contradiction in what we *mean* by know. In other words, if there *were* an illusion here, it could not be one of consciousness. (And that is precisely what we intended to say.)

Now although this discussion might seem a little opaque, its significance is seminal as revealing the problem with mind. With theories of the universe, we are dealing with what is out there. With theories of mind or consciousness, however, we are dealing with the very domain of ourselves.

Of course, the only way we can have theories of the universe (it seems) is *via* this domain of ourselves. But with theories of the universe we do not consider mind *per se*.

Philosophers might query this, protesting that the status of mind begins with Descartes (in the modern era), and continues to the present day, and philosophers *do* consider its capacities. But there is a crucial difference here in wording. For we take Beaumont's point. We are not concerned with the mind and its variegated history. We are concerned with the *construct* mind and its status.

The two examples in which we were misleading – (1) that we lapse into mind talk and its causality without realising the assumptions we have made, and (2) that we make statements about consciousness without realising that the assumption of the powers of consciousness encompass everything and we cannot select convenient particulars exempt from it – illustrate the difficulty of even approaching the topic.

We said that our first and negative demonstration, to show that consciousness as a theory is prescientific, involves scrutinising the various aspects of consciousness. Doing so, however, involves unpicking deep assumptions we never challenge. We do not do so generally because these assumptions are built into the way we speak. But deeper than that, these assumptions in our language arise from our biological nature. It is to that biological *source* that we must penetrate. And we must do so even though apparently we are making those assumptions as we speak. This will require, therefore, a slow and scrupulous approach.

The second part of the negative demonstration, to show why the theory of consciousness arises (as do all theories), results actually from the positive thesis: that of the theory of brain-sign itself.

Brain-sign theory does not begin by trying to accommodate or reconcile physicality with something it does not understand or has failed to define, i.e. consciousness. Indeed brain-sign theory regards the main inhibitor to establishing a science of the brain (a genuine *neuroscience*) as the notion of consciousness itself. Like God, consciousness ‘gets in the way of’ the development of the concepts of science because (as a theory) it pretends to explain something about which it ‘knows’ nothing. Thus brain-sign theory, as a conceptual framework, regards itself as a propaedeutic for neuroscience.

As biological, brain-sign theory assumes the evolutionary process. Indeed, and unlike consciousness (which is unmotivated), it is evolution, and evolutionary advantage, which has allowed brain-sign to come into being.

This account solves one of the great mysteries: what is the relation of other creatures to us? Are *they* conscious? Does consciousness go all the way down to amoeba? If so, how conscious are other creatures by comparison? These mysteries dissolve with brain-sign theory. No creature is conscious, including ourselves.

Brain-sign theory is a monism: we are not two kinds of things, physical and mental. On the other hand, as we shall discover, as a monism, aspects of our being perform different kinds of biological function.

The advantage of the theory as a result (addressing Beaumont’s concern), is that we evade incomprehensible or irreducible kinds of entities as perception, thought, feeling or sensation. Explanation is wholly within the physical universe. On the other hand, the status of

psychology and cognitive science are then under question. For removing entities those disciplines (largely) take as foundational means those disciplines require fundamental review.

Inevitably brain-sign theory impacts our self-understanding, and crucially, our relation to others with whom we inhabit the planet. For example, since in brain-sign theory beliefs do not exist, the notion that beliefs are either justificatory or at least explanatory of our actions is invalidated. This has consequences for the understanding and legislation of human conduct. On the other hand, the *invisible* problem of *how* physical organisms can be in communication with each other (for consciousness and mind are solipsistic presuppositions of our cultural self-explanation) is solved.

Most fundamentally, however, neuroscience is offered a ground upon which to develop as a genuine science. In so doing, it will assume its proper place in the sciences, i.e. as foundational. For neuroscience will establish how any science is possible for us as the organisms we are. The corollary is, however, that the boundaries and limitations of science are established. *Nothing*, it will be determined, can be taken as foundational in any absolute sense – and that includes brain-sign theory itself.

\*

### *Overview*

Before demonstrating the implausibility of consciousness, there is an outline of what consciousness is taken to be. This is Chapter 2. Having done so, we investigate what is wrong with consciousness as a theory, Chapter 3. Chapter 4 goes deeper into the equation of consciousness with knowledge, since it is knowledge that is the scientific problem.

In Chapter 5 we begin to develop the theory of brain-sign *via* its scientific model. This is followed in Chapter 6 by a more extensive account of the function and structure of brain-sign. Chapter 7 is devoted to language, as reconceived under brain-sign theory. In Chapter 8 we investigate the change in human disciplines, including science itself, once consciousness is banished from the organism. Chapter 9 investigates our self-reorientation viewed from the stance of brain-sign theory. In Chapter 10 we conclude with an outline of future work, theoretical and empirical.

# The Theory of Brain-Sign

## 2.0 What is Consciousness?

To address the issues of consciousness, and question its existence, it is necessary first to identify what it has been taken to be. At the extreme, attitudes are polarised between those frustrated by the lack of a convincing answer,<sup>5</sup> and those who consider it too obvious to be considered.<sup>6</sup>

Sigmund Freud, probably the greatest influencer of human self-perception in the twentieth century, now consistently denigrated by more ‘scientific’ psychologists,<sup>7</sup> wrote in his *Introduction to Psycho-Analysis*:

‘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; 14).

This statement, somewhat preceding the current debate, encapsulates the conditions of consciousness. No one knows what it is, but since we have direct acquaintance with it, we are capable of confirming its existence.

Fifty years later, Ned Block said: ‘In the case of thought, we actually have more than one substantive research programme...but in the case of consciousness we have nothing – zilch’ (1994; 211).

Whether substantive research programmes for thought can be of value without constructive ideas about consciousness (since one’s thoughts are deemed to be conscious) is questionable; but even this proponent of consciousness acknowledges that the grounds of his theories are not merely questionable but as yet invisible.

And only yesterday (comparatively), Scott Jordan and Dawn McBride express the chaotic condition of what is termed ‘consciousness studies’. ‘*If* an integrated science [of consciousness] comes to be...’ (emphasis added) (2007; xii).

---

<sup>5</sup> Such may be David Papineau (2003).

<sup>6</sup> Such definitely is Bernard Baars (2004).

<sup>7</sup> Frith, characteristically, says: ‘Freud may have had a big influence on literary criticism...but he was no scientist. He wasn’t interested in evidence. I study psychology scientifically’ (2007; 3).

This is the status of the topic. Nevertheless, the purpose of this chapter is to describe the various properties and characteristics of consciousness (as supposed). We will return to these throughout the subsequent text.

## 2.1 The fundamental premise: knowledge

We begin with a convenient quote from an individual not normally referred to in modern Anglo-American mind literature (though of interest elsewhere), Friedrich Wilhelm Joseph von Schelling. Schelling's position, elements of which may be considered very modern, is that nature is a self-developing teleological process which culminates in its emergent capacity for knowledge of itself as consciousness in man. The following was published in 1800.

'All knowledge is founded upon the coincidence of an objective with a subjective.'<sup>8</sup> For we *know* only what is true; but truth is generally taken to consist in the coincidence of presentations with their objects.

'The intrinsic notion of everything merely *objective* in our knowledge, we may speak of as *nature*. The notion of everything *subjective* is called, on the contrary, the *self*, or the *intelligence*. The two concepts are mutually opposed. The intelligence is initially conceived of as the purely presentative, nature purely as what can be presented; the one as the conscious, the other as the non-conscious. But now in every *knowing* a reciprocal concurrence of the two (the conscious and the intrinsically non-conscious) is necessary; the problem is to explain this concurrence' (1978; 5).

We will not discuss Schelling's theory of *transcendental idealism*, but restate significant elements which are essential for consciousness, and can be detected, in much the same form, before and after him to the present day.

1) There is the division between consciousness and the world. Consciousness is not the world as nature which is the 'non-conscious'. (Schelling has been attributed with inaugurating the notion of the un- or non-consciousness.) In so being, it is a dualist posit from the world (though Schelling proposes to unify the two in his *Naturphilosophie*).

---

<sup>8</sup> Cf. Aristotle: 'Actual knowledge is identical with its object' (1995; 685). The difference between Aristotle's view and Schelling's (as representation) is addressed below.

2) The conceptual division is termed the subjective and the objective. The subjective is the self and the objective is *of* the world object. They are ‘opposed’.

3) However there is a state which is termed knowledge. Knowledge is the coincidence of what is presentative (or is a representation) and what is (re-)presented. This, of course, coincides with the word con-conscious as with-knowledge.

4) Since the *intelligence* or *self* is presentative (i.e. where representations happen), the ascription of consciousness is to the self. By contrast, what is not conscious is nature; its occurrence in terms of consciousness (i.e. as the objective) is to be presented, or as a representation. It will be represented to a self, or, more accurately, in relation to a self as consciousness.

5) Knowledge is taken as that which is true. Truth is the condition of knowledge in the coincidence of the object, nature, with the conscious representation of nature (the objective) to the self.

6) Self and intelligence coincide. This is because intelligence derives from the fact of consciousness – the location where representations happen to the subject as knowledge *and are true*.

7) Knowledge, which consciousness establishes, has a causal function for us as persons (with selfhood), because we know we can act upon our knowledge, and this allows us to act intelligently.

To illustrate Schelling’s currency, here is William Seager: ‘Descartes vision of the mind is the foundation of modern cognitive science. The linchpin of the upstart science is that the mind is in essence a field of representations’ (1999; 4).<sup>9</sup> *Cognitive science*, which aims to be ‘genuinely scientific’, thus passes in nascent form from Descartes through Locke, Hume, Kant, to its most current incarnation. (Schelling derives his philosophic foundations, at the beginning, from Kant *via* Fichte (1970).)

*Representations* are the elements of cognitive science: namely perceptions, as Schelling may be inducing us to visualise; or ideas (or thoughts), which are more generally regarded as cognitive, since they have propositional content; and so on.

---

<sup>9</sup> Seager continues the statement to divide representations into the conscious and non-conscious. We will come back to this.

## 2.2 The I, self or ego

However, for Schelling, and the history before him (though not universally), a key facet concerning representations as *conscious* was their relation to the self, or the I: i.e. that ‘entity’ which *has* intelligence. Cognitive science, as Seager rather restrictively characterises it, has lost the self, and is concerned with representations in a neutered (or information processing) way. This can be seen in the Sloan Foundation’s 1978 commissioned report (unpublished)<sup>10</sup> for the determination of the mission for cognitive science, *viz.*: ‘To discover the representational and computational capacities of the mind and their structural and functional representation in the brain’.<sup>11</sup>

But Descartes said: ‘*I* think, therefore *I* am.’ With Kant the very possibility of consciousness depends upon the I. Here is an example of Kant’s position as stated in the A (1<sup>st</sup>) edition of the *First Critique*:

‘It is only because I ascribe all perceptions to one consciousness (original apperception) that I can say of all perceptions that I am conscious of them’ (1933; 145).

The knowledge statement in the B (2<sup>nd</sup>) edition goes:

‘It must be possible for the “I think” to accompany all my representations; for otherwise something would be represented in me which could not be thought at all, and that is equivalent to saying that the representation would be impossible, or at least would be nothing to me’ (152-153).

For representations to exist, and thence the possibility of *knowledge* (as I can think it), I must be capable of them as *my* consciousness.

Thus for what we might call the formal or original notion of consciousness, the I’s role is foundational. For Kant, the (transcendental) I was *not* a representation, but the condition of the possibility of consciousness. Kant, therefore, considered the I as of an unspecified order of being by which (mental) representation could exist as consciousness –

---

<sup>10</sup> The background to the Sloan Report can be found in a paper by George A. Miller, one of its authors (2003).

<sup>11</sup> Quoted from Bechtel & Graham (1998; 69). They say that cognitive science has developed from this original statement, although the statement remains the case as a significant element of cognitive science. They question whether, in its development, it will break apart as *a* science. Miller, in the paper above referenced, now prefers to speak of the ‘cognitive sciences’.

but not as physicality.<sup>12</sup> If the mind (referred to in the Sloan Report mission statement) is supposed to be predicated on consciousness, then it is difficult to understand what the words of the statement intend by not including the ‘fact’ that the non-physical I is a foundational condition of conscious representations. Indeed, since by definition (as Schelling has it) consciousness is ‘opposed’ (in concept) to a nature which must include the physical brain, how could these representations be ‘in the brain’? Cognitive science, in these terms (i.e. without a solution to the opposition), appears to be either a mistake, or a *non sequitur*.

The evidence of the I of consciousness is that we are aware of ourselves as well as the world around us. It is to us that joys, agonies and accidents happen, and we are as certain of this as we are that we exist. Kant referred to this state as *apperception*. Thus to *be* conscious is to be aware of ourselves *as* conscious, at least in principle. This notion is crucial to Kant for the possibility of knowledge.

The word apperception derives, etymologically, from ad- (ap-) = to, and perception. It means to-perceive. What we perceive, in this usage, is ourselves. In other words, consciousness gives us knowledge, and one of those ‘perceptual’ capacities of knowledge is of our self as *our* self. The *sense* of our self (rather than perception-of, which is metaphorical or allusive) is available to our self by an act of consciousness, which belongs to our self as the *being* of intelligence.

It is obvious that all this would imply an extraordinary achievement of nature (if it is the case, and the universe is purely physical). It is also deeply obscure. For consciousness seems to do a great deal of work, all of which is inexplicable in any other terms than itself.

Little wonder Seager ends his book with these words:

‘In trying to explain consciousness...the standpoint from which the explanation is offered and in terms of which it is understood, contains the very thing we want to explain’ (1999; 251-252).

I.e. there are no other (e.g. reducible) explanatory terms. Consciousness as a theory, Seager says, is ‘tautological’.

---

<sup>12</sup> As Heidegger states it: ‘[Kant] continually inculcated that the ego is not a representation, that it is not a represented object, not a being in the sense of an object, but rather the ground of the possibility of all representing, all perceiving, hence all the perceivedness of beings and thus the ground of all being’ (1982; 128).

### 2.3 Interlude

It is important to remark, at this point, a motivation of both Descartes and Kant. Both wanted to lay the metaphysical foundations for the possibility of science. To do that they had to demonstrate that objective knowledge was possible since, as in Schelling's words, the universe out there (nature) was taken to be the object, while in here was the subjective, not of the universe. Effectively Descartes (a scientific pioneer) did so by God's guarantee that he was not deceived in his perceptions and thoughts. Kant, a profound admirer of Newton, did so by far more elaborate means. But unfortunately he relied upon (scientifically) unexplained transcendental and *a priori* posits and consequent arguments which imposed the mechanisms and criteria of mind upon empirical input, thus extracting the mind from the universe completely (which Schelling later attempted to overcome), and were effectively self-justifying. In other words, both these seminal thinkers in the history of the topic had practical interests in their search, because it was evident to both that scientific success had become a significant fact. However neither succeeded, if we suppose consciousness to be of the physical universe.

### 2.4 Inner and Outer

The modern programme, to demonstrate consciousness as of the brain, makes the assumption that consciousness is in the head. It certainly seems we look out from our eyes, and that behind our eyes is where we are.

Descartes associates the workings of the mind with the workings of the brain in some respects, for the brain is attached to the nerves throughout the body from whence signals signifying e.g. thirst or pain arrive at the brain and are 'transferred' to the mind, providing knowledge of them (Chapter 6 of the *Meditations*, 1968).

The significance of this, for our purpose, is the notion of boundary. Each of us is conscious, and this is bounded by our bodies. We understand our consciousness, and its properties and characteristics, to be identifiable features in our existence as persons; and our existence is determined by the limits of our physical being, which comes into the world and disappears from it. Although medical opinion is not absolute, to be incapable of consciousness effectively means we are dead.

There is, of course, ambiguity about whether our *body* is part of the world while our consciousness is not. ‘Out there’ is the world, as opposed to ‘in here’ of consciousness. In principle, any part of our body could be revealed for others to see as the world: but no one will ever see our consciousness or experience it, despite neuroscientists’ assertions that fMRI scans confirm consciousness (as Frith above). Thus although our body is the boundary of consciousness, it is a boundary with a different emphasis from our skin surface.

The notion of *privacy* links consciousness, or the mind, with its historical antecedent or counterpart, the soul, which defines our unique individuality, and our responsibility for our actions before God at judgement. Both Descartes and Kant intermingle the terms mind and soul, and it is occasionally mentioned today.

There is a fundamental question (to which we have referred) as to whether it can be said of the *brain* that it ‘can see or hear...have experiences...know or believe something,’ as neuroscientists are prone to asserting. This question is raised by Maxwell Bennett and Peter Hacker in their extensive attack on the ‘conceptual confusions’ of neuroscientists (2003; 70). We will address this in other chapters, but at the moment we will rest with the distinction between the brain being in the head, and consciousness being in the head: both are deemed so.

A consequence of this, of course, is whether and how we communicate with others, since our minds, being entirely private, are inaccessible to others. (Indeed, in our worst moments, privacy is terrifying isolation.) This has been a significant conceptual problem in the history of the topic. We merely mention it at this stage.

(There has also been the question of whether, since we can only know our own mental states, anything ‘out there’ exists at all; but since we assume there is an ‘out there’ as well as an ‘in here’, we will not engage that.)

## **2.5 Conscious experience**

The term *experience*, widely used both historically and currently, has become problematic. We say we are conscious, but we also say we experience. ‘(As I’m experiencing it) The sun is very bright today.’ This seems a more likely expression than: ‘(I’m conscious that) The sun is very bright today.’ We shall investigate it directly.

The first thing to say is obvious: it is because we are conscious that we can experience. What is not conscious cannot experience. What does this imply?

We have already identified consciousness as the capacity for knowledge available to the subject. When we refer to our experience, to what do we refer? Presumably knowledge because being conscious we can state that we do experience *knowingly*. But what distinguishes experience from consciousness? The *Chambers Dictionary* (1983) defines experience as: ‘test, trial’; ‘practical acquaintance with’. While consciousness is a more abstract expression of knowledge, i.e. knowledge as an ontological construct, experience has a visceral sense, knowledge as *endured*.

In other words, consciousness is the *mode* by which knowledge exists in the universe, including the fundamental role of the subject, the I. Experience is the *manner* in which knowledge *can* exist for the subject, or I.<sup>13</sup>

Returning to the examples above, when we say: ‘(As I’m experiencing it) The sun is very bright today’, we are noting that, as subjects, we are *undergoing* an experience, enduring the knowledge; we do not say this is because we are conscious, because that is a given for we are all conscious. On the other hand if we had said: ‘(I am conscious that) The sun is very bright today’, we would be referring to a knowledge state of which we are in general capable, and this is the specific one.

What the I of consciousness cannot do is exit consciousness to observe *its* (consciousness’) nature, because it is definitionally an element of the structure of consciousness *as* knowledge. I.e. one cannot experience the *nature* of consciousness, because experience is knowledge as consciousness in its manner of being (endured).

This leads us to the problem of *qualia* or phenomenal properties. Since the topic has been in consciousness studies for decades, we must address it critically here, so discounting it from any further attention.

Here is how the problem arises in the words of John Searle (from the debate with Bennett & Hacker).

---

<sup>13</sup> Bennett and Hacker refer to ‘experience’ as ‘ill-defined’ (2003; 264f). Clarifying it, as we do here, ‘solves’ its ambiguity. For one difficulty they impose upon themselves is their insistence upon the *person* being conscious, rather than the brain. Unfortunately, as John Searle rightly says (in the 2007 debate), they never define persons, which strips credibility from this aspect of their account. ‘They never tell us what a person is, but I think it is crucial to the whole account and indeed to the whole discussion’ (2007; 107).

‘[States like drinking beer or listening to Beethoven’s Ninth Symphony] are *subjective* in the ontological sense that they only exist as experienced by a human or animal subject.... Because of their subjective, qualitative character, these states are sometimes called “qualia”.... “Consciousness” and “qualia” are simply coextensive terms’ (Bennett *et al.*: 2007; 98-99).

The problem is, Searle misconceives the notion of consciousness. It may be correct (if one supposes consciousness exists) that each subject experiences ‘events’ differently. My Beethoven’s Ninth may be different from yours because your and my experience of it involves differences of e.g. background and knowledge of it, and music generally. Given we are both conscious, and consciousness involves a subject foundationally, then (however we express this) what is experienced is different. But that does not mean there is a *qualitative* difference in the experiencing *per se*, a different *qualia*. Nor that that *qualia* is different from drinking beer.

I hear Beethoven’s Ninth. Experienced by me, it is *my* consciousness. I might say: ‘That performance was marvellous’ because I am commenting on the experience, and I may say more about it. But I cannot exit my consciousness to say that the state *itself* has a *qualia*, by contrast with the performance, because being conscious wholly bounds my existence as my experience. My commentary adds nothing on the *experiencing*, for I know nothing of what consciousness itself is.

(It is worth reminding ourselves that Beethoven could not hear his Ninth Symphony. His creation occurred entirely in his head. Nonetheless, we must suppose that, approximately, what we hear is what he intended as he ‘heard’ it.)

Of course I can speak of qualities. They exist *descriptively* because there is consciousness.<sup>14</sup> When speaking of the qualities of the performance, I am speaking of the performance *per se*, not my experience – even if I state that I experienced the performance.

---

<sup>14</sup> Even Locke, who has been accused of originating the idea of *qualia*, actually says: ‘[By perception] we come by those ideas we have of *yellow, white, heat, cold, soft, hard, bitter, sweet*, and all those which we call sensible qualities’ (1997; 109). Locke does not say yellow or heat *is* a sensible quality; he says we *call* it a sensible quality, and it arises from what is conveyed from the object. Later (557) he says it is the appearance of what we *call* the whiteness of the paper that allows him to know that the quality or accident exists *as the object* which caused the idea.

Nor am I speaking about ‘my’ representation of the performance.<sup>15</sup> I cannot exit myself to see that the performance is a representation.

What relates me to the performance *ontologically* is my being *as* conscious. I cannot experience the performance as what goes on in the world (e.g. the compression waves produced by the instruments) because no one can. It is consciousness that allows me (as organism) to describe aspects of the world as having properties and qualities (e.g. being loud). But consciousness does not have qualities; consciousness bestows qualities and properties – it is how the mode of relating to the world and *its* qualities and properties comes into being for a subject. The representation of the performance ‘coincides’ with what is out there, and this (coincidence) is knowledge which, as experienced, facilitates description of what is out there (as the theory goes).

Thus ontologically, experience and consciousness are the same, but usage of the particular term arises upon context. It is not the case, however, that quality and consciousness are the same, because we do not experience quality. Qualities (as reportable) are of the object (or event) we experience.

With a different emphasis, it is wrong to suppose one can experience the redness of the sky (as it seems to me) independently from being conscious of the sky, because experience of the redness is the *manner* of knowledge (properties) of the sky. So David Chalmers must be mistaken in claiming that: ‘The really hard problem of consciousness is the problem of *experience*’ (1995-7; 10). This simply says the hard problem of consciousness is consciousness. Nagel made this seminal mistake in claiming there is a ‘what it is like’ for an experiencer (1974). We cannot exit consciousness to observe our ‘what it is like’ compared to anything. Even if one introduces the notion of *reflection*, as e.g. reflecting upon the experience we are having, we are still simply *being* conscious.

As Schelling says, there is a world out there, and there is consciousness. Subjects do not wander the world experiencing the quality of *things*, then sometimes the qualitative material of consciousness. Subjects are where consciousness happens.

\*

---

<sup>15</sup> As, for example, Michael Tye (1995, 2000).

Now let us consider the status of subjectivity itself. Consciousness is knowledge, and knowledge is the truth. It has been precisely the problem in epistemology that, since consciousness exists for (or as) a subject (the separate and unique individual), how could it achieve objectivity, i.e. knowledge? (Descartes' problem which God solved, and Kant's.) Nonetheless, objectivity about the world *is* the presupposition of consciousness (as a theory). Searle proposes the brain 'makes' consciousness, and does so as the *subjective*. 'These states, qualia, are entirely caused by brain processes' (Bennett *et al.*: 2007; 99). Even were we to accede this, we would still ask why.<sup>16</sup> Why create experience *subjectively*, with the consequent problem of how it could render objectivity?

For now we have the problem of how the brain (as conscious) could be capable of exploring the qualities it has created *as* subjectivity. By claiming consciousness has (or is) *qualia*, there is the supposition that to identify *qualia* consciously, we (the subject of consciousness) are capable of being objective (having an objective representation of our own subjective qualities, thus obtaining knowledge of them) somehow independently of the fact they (the qualities) are still *necessarily* subjective. This appears to involve a complete confusion of suppositions.<sup>17</sup>

This is made plain in a quotation from Joseph Levine.

'Consider the throbbing character of a pain in one's toe after stubbing it, or the way a gleaming fire engine looks in the sunlight. In both cases there is a strong intuition that we are enjoying experiences – conscious, mental states – which can be typed by their qualitative character' (2003; 57).

According to Levine, the existence of the qualitative character of conscious mental states (experiences) derives from his 'strong intuition'. Levine asks: 'What sort of property [is that qualitative character, the quale]?' But again, as with Searle, this hypostatizes mental properties (the throbbing *character* of pain, the *look* of a gleaming fire engine) as if they are really objects in the world which could have properties (though actually they are consciousness itself, apparently). This requires we could be super-conscious (whatever that might be) to 'have' them as objects (but where?) of which we would then be additionally

---

<sup>16</sup> He actually proposes that consciousness offers discriminations which unconsciousness could not, but he does not tell us why (1994).

<sup>17</sup> What might be considered startling is that *prima facie* Searle appears to recognise this in saying: 'I *cannot* observe my own subjectivity, for any observation I might care to make is itself that which was supposed to be observed' (1994; 99). Yet he does not draw the obvious conclusion.

conscious. This must be an ontological debacle. Intuitions, if such exist, are not the route to adequate analysis.

In sum, if the ontology of what consciousness is supposed to be is considered, *qualia* are an absurdity.<sup>18</sup> On the other hand, it is certainly the case that *consciousness* is problematic. How is knowledge with consciousness as its mode possible? How is experience as its manner possible? And how could they be causal?

\*

Let us clarify further aspects of experience. What we do as conscious is see, think, feel, sense, and so on. These are what experiences are. (We do not have an ‘experience of seeing’.)<sup>19</sup> If I see a sunset, I experience the sunset seeingly. If I hear a Beethoven symphony, I experience it hearingly. If the symphony is Tchaikovsky’s Sixth, I hear it and I confer on the hearing tragedy, i.e. it gives me the *describable* sense of being tragic.

Tchaikovsky’s Sixth, as played, is compression waves of the air. But as consciousness it has been converted for me both into hearing as the manner of knowledge, and knowledge in a particular manner as reportable, tragic. But I do not *hear* it tragically. Even if I say it gave *me* pleasure, I am still commenting on my relation to the object of the experience.

Descartes’ description of pain operates the same way. Because I feel pain I can take appropriate action. But I do not feel pain (the throbbing toe) as a painful quality. That it exists (in Descartes’ account) is for me to act on it causally as knowledge. We certainly say we experience pain, but we experience (e.g.) *the toe* painfully. We have no idea what pain is *intrinsically* beyond its knowledge function (though materialists will say it is neurophysiological).

---

<sup>18</sup> This ‘absurdity’ is not the equivalent of Bennett and Hackers’ description of *qualia* as ‘misbegotten’ (2003; 292). The problem for their analysis is that they reject the notion that there is anything to be accounted for. E.g. ‘If I see a red apple, the only object involved is the red apple, not my experience of seeing a red apple’ (286). More will be said later concerning what is involved here.

We note also, to which we shall return, that this account is different from Dennett’s since he explains away *qualia*, referring to them as ‘discriminations’ rather than e.g. colours (1991; 373). At the same time, he retains the disposition of consciousness *per se*, which is extant throughout his text. E.g. how does he *know* colours are discriminations? The question should be: Why are discriminations (if they are) thus?

<sup>19</sup> This kind of phrase confuses the concept of experience, as in John Heil’s words: ‘You might boggle at the thought that your *experience of seeing [a] tomato* could be an occurrence in your brain on the grounds that nothing in your brain (as a technician observes it or as you might observe it in a mirror by an ‘autocerebroscope’) resembles your *experience of seeing the tomato*’ (emphasis added) (2003; 223).

In what way can we say that Tchaikovsky's Sixth generates knowledge? We want to hear it again. We have experienced aurally as knowledge, and can say of it that it was tragic. I may say to you: 'You haven't heard the Sixth? It's wonderful.' And I could not do that unless I had knowledge of it by experiencing it. But I do not want to hear it again because of the qualities of the experience: I want to hear it again because of what it is in my experience *as* it. I want to be *its* being again *as* conscious (though I may not realise, conceptually, that this is what happens).

Almost everything we experience is not turned by us into an explicit description. By so (not) doing, almost everything we experience cannot be described in terms of qualities. But that does not make experience *per se* problematic.

Kant distinguished experience and knowledge because of his empiricist and transcendental model of consciousness. For him, knowledge begins with experience (empirical knowledge, knowledge from the senses). He did not suppose experience was a means to investigate the qualities of consciousness. He would have considered such a notion bizarre.

## 2.6 Illumination

The metaphor of consciousness, which some consider a fact, is 'being in the light', and is distinguished from being in the dark.<sup>20</sup> Of recent times this has become a standard analogy as to whether or not we are so-called zombies. Zombies, those who operate without being in the light, 'all is dark inside', are identical to us but without consciousness (Chalmers: 1996). Why, being conscious, are we in the light and not in the dark? What underpins this distinction? What significance does it have?

In one sense the answer is straightforward. Being in the light means two things. Because we see, we see what is there. It is lit up for us as what is. Secondly it is metaphoric; we see with the light of reason. The light of reason derives directly from the fact that we can see, and is why perception is the first feature of the notion that we know. We shall come back to this, critically, in later chapters. But a few more words are in order here.

---

<sup>20</sup> 'Why doesn't all this information-processing go on in the dark?' (Chalmers: 1995-7; 13)

The connotation of light with reason is associated with the relation of the human to divinity. This comes associatively/historically from the sun as God, the sun being the source of light. God knows all because he made it, and he illuminates it.

The Presocratic fragment 64 in the Diels-Kranz numeration of Heraclitus, translates as: ‘Thunderbolt steers all things’ (Kirk *et al.*: 1983; 197-8). Heidegger’s version is: ‘But lightning steers (in presencing) the totality (of what is present)’ (1975; 72).<sup>21</sup>

In the Bible God creates heaven and earth.

‘The earth was without form and void, and darkness was on the face of the deep; and the spirit of God was moving over the face of the waters. And God said, “Let there be light”; and there was light. And God saw that the light was good.’<sup>22</sup>

By being thence illuminated, Heaven and the dry land called Earth could be *formed*, (etc.).

Light is the light we see, and we see by. It is a divine gift. The first stage of knowledge.<sup>23</sup> For Heraclitus, the association of light is with control of the universe. Heidegger’s interpretation is that, for the early Greeks, lightning (i.e. God) makes what is present in its totality and steers by it; but it is made present *for us* too (as experienced), and allows us to steer intelligently by what is *as* present.

By the time of Plato’s *Republic* and the Cave parable (1974), there is a graphic association not only of light with reason and truth, but also as it inhabits men’s minds – in particular philosophers. Firstly, in the cave, men associate reality with the shadows of their forms cast on the walls by the fire in the cave; then they look at the fire itself as reality, though it pains their eyes; then they are dragged up out of the cave, and by a slow process of painful adaptation they see shadows cast by the sun, reflections in water, men themselves as objects lit up, then the moon and stars; and eventually they look directly at the sun – as simile, at the good, or what is true. Socrates says:

‘But in my opinion, for what it is worth, the final thing to be perceived in the intelligible region, and only perceived with difficulty, is the form of the good: once seen, it is inferred to be responsible for what is right and valuable in anything, producing in the visible region light and the source of light, and being in the intelligible

---

<sup>21</sup> For discussion on the relation to Zeus and earlier notions of the divine sun, see Kirk *et al.* (1983; 198-200).

<sup>22</sup> (Genesis, chap. 1, verses 1-4, RSV)

<sup>23</sup> E.g. Kant’s ‘experience’.

region itself controlling source of truth and intelligence.... It won't be surprising if those who get so far are unwilling to involve themselves in human affairs, and if their minds long to remain in the realm above' (1974; 321).

That 'realm above' of 'their minds' becomes consciousness: from 'visible region' to 'intelligible region'. Consciousness is thus medium, and experience the manner for what we know, and knowledge is the truth. As such we possess quasi-divinity.

## 2.7 Consciousness, the unconscious & non-conscious

But consciousness is not a complete account. We stated at the beginning of the chapter that Freud was probably the greatest influencer of human self-perception in the twentieth century. Though Freud was not the first to countenance the non-conscious (cf. Schelling above),<sup>24</sup> or its capacity for causing intelligent behaviour, he made the notion a cultural commonplace, and undermined the supposition that what we (e.g.) do necessarily arises from our self-illuminated intelligible reasons.

These days, *most* intelligent activity is understood to be unconscious, or non-conscious. Here is 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; ix).

Although intelligent non-conscious processes may have seemed strange in Freud's time, in the interim, computers founded on the notion of processing symbolic representation under syntactical rules, and neural networks (or connectionism) activated without symbolic representation, have demonstrated that 'mere matter' can effect intelligent work.

But this does not address the topic fully. For computer processing and neural networks are *not* non-conscious but 'in their essential features very like the more familiar

---

<sup>24</sup> This was outlined (e.g.) in the classic study by Lancelot Law White (1960).

mental processes of perception, inference and cognition'. In an engineering (i.e. practical implementation) sense we have no idea what *conscious* mental processes ('of perception, inference and cognition') are. While we may suppose that somehow the brain acts intelligently, we have no idea how the brain as *consciousness* (if it is) acts at all. Therefore the supposition that non-conscious processes are in their 'essential features' like conscious processes is doubly perplexing. Nor, by default, can non-conscious processes be regarded as *mental*.

If on a battlefield a soldier encounters an enemy soldier and must kill him to survive, yet is struck by the realisation the enemy is wounded and cannot defend himself, so is torn by pity and hesitancy; then afterwards and for the rest of his life is tortured by images and guilt, for he did shoot him, could not stop himself...these are the actions, thoughts and feelings, not to mention perceptions, of a conscious human being. It may be, as is commonly stated, that 'there is ample evidence that human phenomena we would like to label "smart" are not always accessible to, or have their origins, in conscious contents or procedures [and they] range from the most basic perceptual processes to the learning of complex systems and the production of creative ideas' (Allen & Reber: 1998; 314). But how could the soldier's *mental* activity exist *per se* were he not conscious, whatever non-conscious background to his existence there was?

At the same time, however, if non-conscious processes are so 'smart', why does consciousness exist at all? What job does it do beyond non-conscious activity?

Antonio Damasio gives a standard answer. 'The drama of the human condition comes from consciousness because it concerns knowledge obtained in a bargain that none of us struck: the cost of a better existence is the loss of innocence about that very existence' (1999; 316). But this is not progress, for it is tautological as we have encountered before. Conscious is knowledge, and we know because we are conscious.

Concerning the *relation* of non-conscious to conscious activity, Edelman and Tononi state: 'In action as...in perception, it appears as if only the last levels of control or...analysis are available to consciousness, while everything else proceeds automatically [i.e. non-consciously]' (2000; 58).

But what *decides* what should be conscious? Alternatively, how does it *become* conscious? And what does '*available* to consciousness' mean (as opposed to *being*

conscious)? The statement is fundamentally confused. Consciousness appears to involve some kind of extraordinary *functioning* differentiable from non-conscious functioning: but what kind?

Later in their book, Edelman and Tononi state:

‘Conscious thought is a set of relations with a meaning that goes beyond just energy or matter (although it involves both)... There are no completely separate domains of matter and mind...but obviously, there is a realm created by the physical order...in which meaning is consciously made’ (2000; 219).

What was confusion has now lost any coherent interpretation. For while proposing to *explain* consciousness and its difference from the non-conscious or the simply neural, the authors are dragged into a twilight zone in which mind both is and is not matter. One senses that what is at issue has not been identified.

John Heil expresses the problem succinctly: ‘Even if we knew that neural tissue arranged in a particular way yielded a feeling of pain, the reason this arrangement yields pain rather than some other feeling (or no feeling at all) remains an utter mystery’ (2003; 235). But expressing the problem is not solving it.

One could reasonably say that the ‘invention’ of the unconscious has not made the notion of consciousness more intelligible. If anything, it has made the situation worse. For as well as failing to explain consciousness, it has made questionable the need for consciousness at all.

## **2.8 Intentionality**

The reintroduction of the notion of intentionality from its medieval origins begins with Brentano (1973). In a straightforward sense it identifies the character of mental states, which was Brentano’s purpose, and distinguishes them from physical states. If I think of Paris, I am thinking *about* the City of Light in France. How could the brain do that, given it is a purely physical object?

But the situation is more complex. If I see Paris from the top of the Eiffel Tower, for example, am I seeing the inner representation of Paris in my mind, or am I seeing Paris itself?

Brentano addressed the notion of intentionality in the empiricist tradition from Locke, and so supposed that intentionality referred to the inner image in its inner domain of the mind. Cognitive science, as discussed, generally sticks with this interpretation: the mental world is the domain of (inner) representations. In our relation to the world, therefore, these representations are the (mental) intermediary with the real world. (This, it is supposed by some, specifically Jerry Fodor, could make them tractable to science, cf. the Sloan report above.)<sup>25</sup>

But Brentano misunderstood the notion of intentionality historically, so his account, appealing back to Aristotle, is muddled. In our context, it is worth outlining this briefly.

Brentano distinguished the mental act from its content, rather than the empiricist notion that simplified mental states as associative reactions (Hume's impressions, Kant's affected mind of experience). He proposed mental acts reach out to the real objects themselves, or are intentional. This recalled the medieval notion of *esse intentionale*. But the medieval notion, which concerned the *being* of the knowing state, was intended to clarify Aristotle's account of the knower 'obtaining the form'<sup>26</sup> of the object known, as (as an analogy) a hand surrounding an object.<sup>27</sup> Thus the medieval notion regards the intentional object as, as it were, the transparent sign of knowledge the knower has of the real object. It *is* a mediating state, but nonetheless a state in which the real object is known directly.<sup>28</sup> Although this notion might be understood to give rise to the modern version of ideas existing in their inner mental domain, it has a quite different intent.

For example, the epistemological problem, as referred to, is how the inner mental state can *correspond* to the external object for knowledge to be obtained in its truth. In its pre-Cartesian form this problem does not arise since the account supposes a different criterion for obtaining knowledge by the mental act. Thus when Husserl developed Brentano's notions (reconstructing them originally, rather than referring back to Scholastic or Aristotelian models) as *Phenomenology*, he surmounted the problems of the subjective intermediary position by what we might term direct (or directed) knowledge. When he states in *Ideas I*: 'The spatial physical thing which we see is, with all its transcendence, still

---

<sup>25</sup> This achieved by 'reducing' inner representations to the symbols of a notional language, which is then interpreted as computable. The implausibility of this is discussed by Jean-Michael Roy (1999).

<sup>26</sup> 'About all perception...a sense is what has the power of receiving into itself the sensible form of things without the matter' (1995; 674).

<sup>27</sup> 'The soul is analogous to the hand; for as the hand is the tool of tools, so thought is the form of forms and sense is the form of sensible things' (1995; 686).

<sup>28</sup> Discussion can be found in Knudsen (1982; 490 and below).

something perceived, given “in person” in the manner peculiar to consciousness’ (1982; 92), he is expressing a different conception of what perception effects (as intentional) to that of the impressed or affected inner domain of Hume or Kant.

It is questionable, therefore, whether the term *intentionality* is appropriately used for the model of consciousness which derives from empiricism, since in that case what consciousness is ‘consciousness of’ is not as ‘directed to’ but ‘impressed upon’, and so problematically isolated from the world.

However, Husserl’s ambition, to obtain apodictic certainty which could underpin all the sciences, rendering (phenomenological) philosophy the science of sciences, was precisely that of Descartes and Kant.

Husserl initiated what became termed the Continental tradition.<sup>29</sup> Heidegger, for example, discarded the notion of consciousness altogether, reacting against the ‘transcendental’ Husserl of the *Ideas*, though he worked still in phenomenological and intentional conceptualisation in his most celebrated opus, *Being and Time* (1962).

We shall return to this in later chapters.

## 2.9 Memory

Memory has a special status concerning consciousness, and when critically addressed in later chapters, the necessity for its reconstruction will become apparent.

The *mind* notion of memories is that we all have them. Indeed our lives are dominated by the fact that we have a history of experiences, all of which are stored away and can return to us either unbidden, at will, or as triggered by some current event. This can be characterised in supposing memory is a great reservoir of our former life. As a person, we are capable of recalling and discussing how our past has been. Indeed, it is almost inconceivable that one could refer to a person without their ability to recall their lives and the significance of their memories of it.

But there is a problem here akin to consciousness itself. Where and how are memories retained? Our experience as consciousness is of the past as present again, sometimes with an

---

<sup>29</sup> ‘Continental’ is a portmanteau word embracing different positions.

almost hallucinatory immediacy. But does that mean, as would be colloquially supposed, there *is* a great reservoir of our former life in our minds? This was Hume's notion distinguishing *impressions*, original encounters with the world, and *ideas*, their more faint re-enactments (1962). But the notion of mind, in this case, is intrinsically problematic, since now we must search not only for consciousness 'in the brain', but for where and how all these memories exist, and in what form.

What is the function of memory? It must be to guide present and future action. The brain, we suppose, is an adaptational organ, and in encountering the world it alters according to the good or bad effect of its actions.<sup>30</sup> How it judges the effect is, of course, an important question. We will not address that yet.

But over and above the brain's adaptational altering (non-consciously, as it were), consciousness is added. In other words, as we have asked before, why should there need to be consciousness which gives us subsequently conscious memories?

(The following refers to what is termed explicit or declarative memory, as opposed to implicit or procedural memory.)<sup>31</sup>

After the original moment of the experience, it is generally supposed to be stored unconsciously and can reoccur. But how can conscious experiences be stored unconsciously? This appears to be a contradiction. We must suppose that from being originally conscious to becoming unconscious before becoming conscious again, there is a method by which the brain can make its contents conscious independently of neural implementation. For at the moment the brain is (e.g.) affected by an external scene it is both neurally altered *and* conscious. But whilst the consciousness of it disappears, the neural alteration remains. At some subsequent time, the neural alteration (as *engram* or *trace*) is reactivated to become conscious again.

There appears to be two things at issue here. (1) Consciousness is knowledge, so whatever is conscious is knowledge. But in addition, whatever is conscious is present. So one could say consciousness is present knowledge, and that could apply to what is now remembered. But whilst original knowledge can be banished to the unconscious or non-conscious before it becomes present knowledge again – and unconscious *knowledge* is, as it

---

<sup>30</sup> Edelman and Tononi (2000) regard this as directly analogous to Darwinian selection.

<sup>31</sup> The distinction is attributed to both: Cohen & Squire (1980) and Schacter & Graf (1986).

were, metaphorical – if the unconscious is purely neural (i.e. not a great reservoir), (2) what has happened to the knowledge as in the recallable memory? For example, where has the redness gone which we can recall as red? Alternatively, when we are conscious of that memory, how has the redness returned?

This leads to another question concerning remembering. What causes it? We can acknowledge that memories can be triggered associatively, and we also try to remember by making present something known in memory. But in addition we are continually bombarded by memories unbidden. To ask what causes them is to ask a fundamental question about brain functioning.

The experiments of Wilder Penfield (1958) demonstrated that electrical stimulation of areas of the cortex generated in the subject comprehensive images, speech, sensations, apparently as quasi-memories. These were evoked directly by Penfield's stimulation: but what causes them when there is no external stimulation? Why are we (in) a stream of consciousness? It appears, *prima facie*, that knowledge, as memory, wants to be present under its own volition. That must be very significant concerning what it is to be a human being.

## 2.10 Imagination

Moreover, we are bombarded by images and thoughts which neither result from present stimulation, nor appear to be memories. These are imaginative constructs. The brain continually re-works material to generate new forms.

Is it correct to suppose imaginative activity is knowledge? In the classic phrase, it can be termed speculative knowledge. It seems remarkable we exist not merely in the immediate, and with memory, but in the projection of the (im)possible. The situation can be well précised from the writings of Gaston Bachelard.

‘Creative imagination has quite different functions from reproductive imagination. To the former belongs a *function of unreality* which has as much psychological utility as the *function of reality*’ (1971; 13)

‘Imagination is always considered to be the faculty of *forming* images. But it is rather the faculty of *deforming* images offered by perception, of freeing ourselves from the immediate images’ (19).

‘A dialectic of values actuates the *imagination* of qualities. To imagine a quality is to give it a value which goes beyond or contradicts the sensible value, the real value.... We seek the *other* in the midst of the *identical*’ (83).

Imagination is the mode in which alternative ways of viewing the world and ourselves occurs. This can be a mixed blessing, for while novel ways may be highly creative, they can also be paranoid and frightening, symptomatic of mental illness.

\*

Kant, who gave imagination philosophical legitimacy, did so not by assigning to it the ‘mere’ role of generating new forms, but as fundamental in the synthesis of the content of conscious states themselves. As he says in a footnote: ‘Psychologists have hitherto failed to realise that imagination is a necessary ingredient of perception itself.... Something more than the mere receptivity of impressions’ – this is aimed at Hume – ‘is undoubtedly required’ (1933; 114). His expression of the (complex process of) synthesis changed between editions as his account, moving from the subjective (empirical) to the objective (the A edition), became the purely objective (the transcendental B edition).<sup>32</sup>

However at the heart of this is a profound uncertainty. Kant may have referred to imagination in the A edition as the ‘blind but indispensable function of the soul’ that engineers the synthesis, and in the B edition should have replaced this with ‘Synthesis in general, as we shall hereafter see, is the mere result of the power of the imagination, a function of the understanding’.<sup>33</sup> But although there is a demarcation between what Bachelard refers to as *creative imagination* and *reproductive imagination*, both are founded upon a mechanism which, with Kant, though subtly worked out, is both conjectural and opaque. The ‘blind...function...’ of imagination does not explain *why* imagination does what it does with the occurring result (which presumably accounts for Kant calling it ‘blind’: and, moreover, he took the result – as he conceived it – for granted).

---

<sup>32</sup> The syntheses are Apprehension, Reproduction and Recognition, and are detailed in the Transcendental Deduction. Husserl developed this, for example in terms of the *inner horizon* by which perception entails seeing, not a flat surface, but a complete entity, though not in immediacy (1982; 84 and below).

<sup>33</sup> Reference here is to commentary by Rudolf Makkreel (1990; 29). Although Kant reproduces the phrase ‘blind...function...’ in the B edition, Makkreel regards this as an oversight on Kant’s part.

Thus, for example, where is to be found the crucial distinction between what is real (e.g. present and visible before us) and what is unreal (e.g. imagined and bearing no relation to what is present) *via* the mechanism (imagination) creating us as consciousness? More precisely, although consciousness is knowledge, and knowledge is the truth, have we, from Kant and subsequent writers, any *scientific* determination of what constitutes reality (as knowledge) as produced in us, *and how?* I.e. the *fact* of imagination (assuming it is a fact) clouds our confidence in consciousness (so created) as offering reality/truth *per se*. At the extreme this is evidenced in what is now known as ‘false memory syndrome’, where we supposedly remember (in the present) what is in fact imagined.

In the *Prolegomena*, Kant states:

‘The imagination may perhaps be forgiven for occasional vagaries and for not keeping within the limits of experience, since it gains life and vigor by such flights and since it is always easier to moderate its boldness than to stimulate its languor. But the understanding which ought to *think* can never be forgiven for indulging in vagaries; for we depend upon it alone for assistance to set bounds, when necessary, to the vagaries of the imagination’ (1971; 59).

Kant’s move from imagination as (perhaps) an independent faculty (A) to the understanding of which imagination is a ‘mere’ function (B), demonstrates his concern that understanding should be the (transcendental) mechanism ‘guaranteeing’ truth (cf. §2.3),<sup>34</sup> rather than a ‘blind...function of the soul’.<sup>35</sup> But this seems not only either arbitrary or derived from rationalistic prejudice, it fails to offer a scientific explanation of imagination and consciousness *per se* – which, of course, with Kant, was inevitable since he intended to preserve the mind from scientific determinism.

## 2.11 Feeling, sensation and mood

We have already referred to two of these as an aspect of knowledge. The one remaining is mood. It is evident that they all are differentiable from perception and thought, though as features of mental states rather than separable items. Mood, as is well known, can be changed

---

<sup>34</sup> By imposing the mechanisms and criteria of the mind on empirical input.

<sup>35</sup> According to Heidegger, Kant ‘shrank back’ from this ‘transcendental power of imagination’ ‘as a third basic faculty, along with sensibility and understanding’ (1990; 110-111).

by alcohol or other drugs, or physical stimulation as with exercise. Is it the case that mood can be defined, as an aspect of consciousness, as knowledge, or is it an exception? The answer would clearly be that it is knowledge, but in terms of a *general* state of the person. ‘I am in a good mood.’

## 2.12 Free Will

The last topic is highly significant. According to Robert Kane:

‘The problem of free will and necessity (or determinism) is perhaps the most voluminously debated of all philosophical problems’ (2002; 4)

The debate is unresolved; could be said to be in a stalemate. However, books continue to be written; one edited by Gary Watson in 2003. But his earlier statement (1995) could not be more indicative of the status. ‘While subsequent developments have sharpened and focused the issues, the basic questions remain what they were in the seventeenth century’ (181).<sup>36</sup> This is highly relevant, for consciousness was a creature of the seventeenth century.

Two fundamental reasons underlie our notion we have free will. They result, as in previous discussions, from our experience, and the historical development of the theory of consciousness. Historically, as mentioned, our freedom as soul is endowed by God, but our resulting conduct in the use of our free will will be judged and our eternal fate thence decided.

In terms of our experience, we appear to be able to consider our options, to make up our minds, to decide. And this is: (1) On the basis of our knowledge; and (2) As our inclination directs us. Although the divine judgement position can be regarded as erroneous, our experience still gives us the sense that our freedom is, for us, vitally important.

Much of the debate revolves around how the debate is to be conducted, and what the notion of free will entails. Thus although terms are established, for example that we are free if ‘we could have done otherwise’, these terms themselves come under attack.

For our purposes, the significance will turn on whether our *sense* of free will is valid, not whether we could in fact be free – whatever that may mean. For as we have noted from

---

<sup>36</sup> Also Wegner (2002); Gazzaniga (2005).

the beginning of this chapter, consciousness in its intrinsic nature has been considered not to be part of the physical universe, and thus is exempt from its deterministic character. That was Descartes' supposition, and Kant's aim. But from the beginning of its history, whether consciousness *could* be outside the physical universe was questioned.

### 2.13 Summary

This brief survey of consciousness has revealed critical indeterminacies. Consciousness is knowledge; it is the mechanism by which knowledge comes into being. But there are fundamental differences in how it is proposed that mechanism (e.g. intentionality) be understood, and perhaps therefore what knowledge entails.

Historically it has been supposed that consciousness does not belong to the physical universe. There has been a parallel between a God's-eye view of knowledge and that of humans, and this is associated with the divine dispensation of knowledge itself. The aim, to render consciousness of this world, indeed in the brain, sets aside its original conception.

On the other hand, in recent years there has developed a confusion about whether consciousness *is* the effective means by which the brain operates (i.e. versus the unconscious). This has placed consciousness under a question mark, for not only is there no agreed view, there are no definitive terms for addressing the issue.

There is also confusion about whether consciousness is actually the provider of knowledge, or is rather a repository of a scientifically unreachable kind of existent (i.e. *qualia*) belonging to each individual, based on the 'fact' of their individuality.

So even before beginning a critique of consciousness, we find no stable account against which to apply scientific theorising. Yet we have stayed entirely within the mainstream, historically and currently. It is worth recalling the words of Thomas Kuhn.

“Normal science” means research based upon one or more scientific achievements, achievements that some particular scientific community acknowledges for a time as supplying the foundation for its further practice’ (1996;10)<sup>37</sup>

---

<sup>37</sup> There is an industry rendering consciousness physical, and therefore presumably 'scientific'. It includes emergentism, reductive physicalism, non-reductive physicalism, supervenience. Failing that, there is

Although consciousness is regarded as necessary for providing the foundational possibility of science, in truth there is no agreed definition, that would constitute 'normal science', of what consciousness is or what it does. And that appears to make science itself problematic.

---

epiphenomenalism, rendering conscious causally inert, and eliminativism that either denies consciousness exists or wholly reduces it. We regard this industry as a side-line to our concern.

## The Theory of Brain-Sign

### 3.0 What is wrong with consciousness?

Consciousness is a theory. It was invented in the seventeenth century. Is it a valid theory? How would we judge its validity?

There are three approaches. The first is to question it theoretically: e.g. are its premises or structure plausible in themselves? (In describing consciousness we have already encountered difficulties in stating what it is.) The second is to question its scientific appropriateness. The third is to consider its empirical plausibility.

Hindering such questioning is a cultural supposition. There is an (at least implicit) assumption that consciousness *must* be a valid theory, else how could we judge whether consciousness is a valid theory? For it is consciousness that would give us knowledge that consciousness is not a valid theory. We have encountered this tautological tangle already.

We do not offer a knock-down argument that consciousness is a wrong theory. We shall accumulate evidence of its flaws, evidence that does not undermine our alternative theory.<sup>38</sup> Indeed, the reason there is no knock-down argument may well arise from the theory of the brain phenomenon we shall subsequently propose.

In this chapter, the first section (3.1.x) addresses consciousness as a theory, finding it endemically confused. The second (3.2.x) identifies empirical factors which undermine the theoretical suppositions, rendering them questionable as science. The third section (3.3.x) focuses on empirical material, demonstrating that, as a theory, consciousness does not sustain the pressure of evidence, and that practitioners have not yet faced up to this.

### 3.1 Representation and correspondence

Central to the topic is representation. And crucial to the topic of representation is correspondence to what is out there which will render it true as knowledge. It seems,

---

<sup>38</sup> In an attempt to address this, such statements are made, even by physicalists as Paul Churchland. 'The medieval and premodern attempts [at Biology]...were downright pitiful, as were the medical practices that were based on them. But why should we expect our understanding of the nature of Cognition...to be any *less* pitiful, prior to making comparable achievements [as modern Biology] in penetrating the structure and activities of the brain?' (2007; 238). However, as mentioned, Churchland still supposes consciousness exists.

superficially, straightforward. Because the mind cannot directly manipulate the world, nor indeed be in direct contact with the world, it performs its causal function by representing the world, and acting on its own representations. But the topic is fraught with problems.

### 3.1.1 Representation from Descartes to the Empiricists

The problems date back to the origins of the modern notion of mind. A key one lies in the criticisms by Arnauld of Malebranche, both influenced by Descartes, and both finding in Descartes' writings grounds for their position.

Descartes was an idealist (in Kant's words, a 'problematic idealist'), and an innate-ist: he supposed everything in the mind pre-exists. 'Seeing' a tree (perceptual cognition) results from the impact of the corpuscular rays from the tree upon the retina, thence entering, by nerve transmission, the 'corporeal imagination' (imagination in the brain) and triggering the idea of 'tree' in the mind (perception and cognition). It was impossible, for Descartes, that physical processes could cause mental processes which were spiritual. The 'triggering' was therefore not causation, but the idea-capacity 'reading' directly the natural signing in the corporeal imagination. He says in *Comments on a Certain Broadsheet*:

'There is nothing in our ideas which is not innate to the mind or the faculty of thinking, with the sole exception of those circumstances which relate to experience, such as the fact that we judge that this or that idea which we now have immediately before our mind refers to a certain thing situated outside us. We make such a judgement not because they transmit the ideas to our mind through the sense organs, but because they transmit something which, at exactly the right moment, gives the mind occasion to form these ideas by means of the faculty innate to it' (1985; 304).<sup>39</sup>

It was Descartes' view that what was 'out there' was not seen *per se* by the mind. For example, light and colours as we see them, for Descartes, are not in the mechanical universe. So how could it be that the requisite ideas were already in the mind? They are placed there by God. (In other words, although Descartes does not directly parallel the illumination of the world by the sun with the illumination of the human mind as had Plato, he attributed the

---

<sup>39</sup> Reference in the following description is to the Introduction by Stephen Gaukroger to his translation of Arnauld's *On True and False Ideas* (1990).

*illumination* of the mind to God by forgetting, as it were, the antecedent association of the sun, and its illumination, *with* God.)

However, Descartes also gave a different account (in *Le Monde*) from the direct (or semantic) one stated above. In this version, responding to the position of faculty psychologists, he appears to introduce an additional step such that the mind *interprets* the sign of the natural process, for example the corporeal sign of the tree from the corpuscular rays on the retina, and then has the mind *represent* that interpretation to us. In other words, he seems to interpose the mind as representing to us the idea of a tree, rather than (us as) the mind perceiving the tree directly. What is thereby introduced is us as *spectator* of representations.

‘It is our mind that represents to us the idea of light each time our eye is affected by the action which signifies it’ (1985; 81).

As Gaukroger says: ‘The semantic [or direct] interpretation has more claim than the other to be taken as Descartes’ considered view of the matter...[but] it cannot be said confidently that he ever finally and unequivocally accepts or even fully understands it’ (1990; 26).<sup>40</sup>

Considering the previous chapter, we see the emergence of the two approaches to the perceptual act, and indeed the two approaches to mind. In simplified terms, the first has the perceptual act as a direct grasp by the mind of the object, and in the second there is introduced the representational fabric as the mind’s creation, the representations being what we (as spectators) see.

Arnauld’s criticism of Malebranche (ignoring the divine aspects, to which we shall return briefly in the next chapter) is that the latter proposes we do not see objects directly, but rather intelligible objects, i.e. those the mind presents to us. And of Arnauld’s various criticisms, the relevant one here is that Malebranche’s model (1980), as Arnauld interprets it, is based upon the comparison of ‘what we see’ with pictures, corporeal things, i.e. things of nature which represent something other. Arnauld’s position, however, is that mental seeing (as a spiritual act) is of an entirely different ontological order from the kind of representation that pictures achieve; his example of the latter is actually the image in a *camera obscura*. For the nature of the corporeal representation is not as being seen; a subject must exist for it to become seen, and this is entirely different from the condition of the mental process of seeing.

---

<sup>40</sup> Gaukroger refers here to John Yolton’s account in *Perceptual Acquaintance* (1984).

As Arnauld says in his criticism of Malebranche:

‘It is no longer *the perception of bodies* that he [Malebranche] calls an *idea*, but rather a particular *representation* of bodies which he claims is needed to make good sense the absence of any body that can be joined intimately to the soul, so that the *representation* is thereby *the immediate object and what is closest to the mind when it perceives it*. He does not say that it is in the mind or that it is a modification of it, which he would say if he had meant by it only the perception of the object, but only that it is “closest to the mind”, because he regards this representation as being actually distinct from the mind as well as from the object’ (1990; 63).

In other words, the word *perception*, which for Arnauld means a direct seeing of the object in the world, has been replaced by the intrusion of representations, and it is these we see. So now in Malebranche’s account as a model we have: real object, its engineered representation and the seer of the representation. The Arnauld model is: real object, the representation of the object as seen.

Note that Arnauld refers to the situation in which the real body is absent, which Malebranche has used as a reason for his representational theory. But Arnauld’s point is that Malebranche confuses seeing with the eye, which in theory would require the object to be present, with seeing with the mind, which in fact must *always* be the case whether the body/object is there or not. I.e. seeing is a mental act; seeing a corporeal representation involves a seer of the representation, but this adds the divide between the seer and the representation, and that divide is what Malebranche has introduced by creating the internal representation as if it were an external representation.

It is evident what is the problem with Malebranche’s theory, as Arnauld considers it. Representations (‘chimeras’, Arnauld calls them (1990; 64)) have been cut loose from both the world and the mind. They are a pure (theoretical) fiction. Of these chimeras Arnauld states critically:

‘One must take as true only what is clear and evident: and one must not make use of alleged entities [representations] of which we have no clear and distinct idea in order to explain the effects of nature’ (1990; 65).

But what actually happens, historically? As in the previous chapter, when the mind model passes to the empiricists, this notion of inner representation has assumed its fully ‘fictional’ (or chimerical) status. With Locke we have:

‘When I say the senses convey into the mind, I mean, they from external objects convey into the mind what produces there those *perceptions*’ (1997; 110).

‘Perceptions’, as Locke uses the term, is *of* ideas, or representations to a seer, not representations of world objects *as seen*. Note Locke’s use of the word ‘produces’.

With Hume we have

‘Those perceptions which enter with most force and violence, we name *impressions*; and under this name I comprehend all our sensations, passions, and emotions, as they make their first appearance in the soul’ (1962; 45)....

Note the words ‘make their first appearance’.

And with Kant we have:

‘The capacity (receptivity) for receiving representations through the mode in which we are affected by objects, is entitled *sensibility*. Objects are *given* to us by means of this sensibility, and it alone yields us *intuitions*; they are *thought* through the understanding, and from the understanding arise *concepts*’ (1933; 65)....

Note the ‘given to us’.

But *what* are these *perceptions* conveyed into the mind (Locke)? What are these *impressions* which can appear in the soul (Hume)? What are these *representations* by which we are *given* objects, etc. (Kant)? And who is this *us* that can receive or perceive these *given* mental objects?

None are explained in any terms other than themselves. They are asserted as self-evident. But if we have no explanation of what a mind or soul or sensibility *is* that can receive or perceive objects from the world, what confronts us is description with neither ontological nor scientific foundation.

As to causality, *how* can a mind have perceptions conveyed (Locke)?<sup>41</sup> How can perceptions enter with force and violence so they appear in the soul (Hume)? How are representations the result of sensibility being affected by objects (Kant)? What mechanisms are being assumed? for *assumptions* they are.

Indeed, these non-explanations are worse than Descartes', for Descartes at least attempted to account for the mind, however implausible his explanation might be. Minds were slices of *quasi*-divinity, though once existing, there is no direct divine intervention. Their content could not be caused by physical processes. But the empiricists invented the media, the causality and the representational state *ex nihilo* that do the mind work. The empiricists kept the language that precedes them, but de-divinised it.

### 3.1.2 Locke's absence of justification

Where have we arrived? There is the real world out there; in here is a substitute world, a world of replicas, and by seeing these replicas (somehow), we have knowledge of the world out there. The inner world is, of course (it is claimed), just as real as the outer world. Richard Rorty referred to it critically as the 'mirror of nature' (1980).

Let us take this further. Locke offers no ontological account of perceptions, i.e. an account grounded in understood physicality or causality. About the *being* of perceptions Locke is agnostic. 'The more probable opinion is, that this [perceptual] consciousness is annexed to, and the affection of one individual immaterial substance. But let men, according to their divers hypotheses, resolve of that as they please' (1997; 311). Were we to say to him, 'But look, this is all terribly obscure. Not one term is explained', we will find his response in his text. '*What perception is*, everyone will know better by reflecting on what he does himself, when he sees, hears, feels, *etc.* or thinks, than by any discourse of mine' (1997; 142) (original emphasis).

In other words, Locke appeals to our experience which he takes for granted (1997; II.i.§19). (Cf. Freud & Block.) He assumes a model 'we all know' because it is self-evident. In supposing we can accurately interpret our experience by reflection, he appeals to a technical capability we must have (*as* mind) to make good the mind's self-revelation; and that

---

<sup>41</sup> Locke was cognisant of the difficulty. In the fourth edition he drew an analogy with Newtonian gravitation. This was not, however, an effective resolution. See endnote 8, Locke (1997; 756)

the *we* to which this self-revelation is apparent *is* the observer of the operations of our own minds. But obviously, by this, Locke does no *fundamental* explanatory work at all. (Perhaps his emphasis of the ontological phrase ‘*What perception is*’ indicates his unacknowledged embarrassment about this – i.e. unjustified assertion.)

Is it possible we can investigate our own minds to identify them *as* minds by self-inspection? This question will be the ultimate destination of our critique of mind, and is addressed in the next chapter. We are making a series of steps towards it.

### 3.1.3 The problem of covariance and misrepresentation

Locke realised that inner representations could not be ontological copies of what is out there. The word ‘tree’ is nothing like a tree, and the image of a tree we see is nothing like the tree as existent. The image (the representation) is not made of wood, nor has leaves, nor is it whatever it is in the world that causes us to see green, and so on. They are *ideas* (as Locke termed them), and are to be distinguished from the external bodies that have caused them.

Yet Locke was inherently ambiguous about this division in proposing *primary* and *secondary* qualities. Primary qualities are those that are ‘utterly inseparable from the [external] body’ (135) which he lists as e.g. ‘solidity, extension, figure, motion, or rest, and number’ (*ibid.*). Secondary qualities are those belonging entirely to the mind, e.g. ‘colours, sounds, tastes’ (*ibid.*), though they are caused by the presence of the external body. Locke considered primary quality ideas as ‘resemblances [of the bodies]’ which have induced them (136). But two problems about Locke’s resemblances are:

(1) What are ideas/perceptions (ontologically) such that they could *be* resemblances? (A painting occurs on a canvas, but there is no canvas in the brain.) Locke, of course, does not tell us. And...

(2) The implication is that if a body is in front of us, we must see it *as* a resemblance. I.e. he proposes (implicitly) that primary quality ideas *covary* with what in the world has caused them. We can rely upon representations to ‘give’ us the world because, for example, when I see a tree, the image I have is caused by this particular tree I see, and *truly* resembles it.

The second problem, as Robert Cummins dwells upon it (1989, 1996), is that covariance as a causal mechanism does not guarantee that what we see ‘gives’ us what is in the world, because we can suffer *misrepresentation* (i.e. correspondence for knowledge as *truth* fails).<sup>42</sup> By this is not meant that, for example, we see a stick as bent when partially in water, for this is regarded as a condition of the world as it appears, or as light conveys the actual circumstances. What is meant is that sometimes we might see a tree as a lamppost. Of course, we might suppose that we have been deceived because the light plays tricks, or the local environment misled us: but these explanations are not adequate. For if, as Locke proposes, we *should* see a tree, given its existence before us and given the notion of covariance, and we mistake it for a lamppost, then covariance *as a mechanism* appears to have failed as an account of how mental states occur in relation to the world. Because Locke has representations as ‘internally viewed’, there is either a radical disconnection between our representations and the world (for it was a tree there, not a lamppost), or a failure of the model to explain the causal process involved in seeing (for the representation of the tree failed to be the object we see). And this applies also to our claiming that we see a lamppost – i.e. our verbal report. In other words, the possibility of misrepresentation seems to put Locke’s account (the model of perception) into question.

### 3.1.4 Slezak’s critique

The problem of misrepresentation has been critiqued by Peter Slezak in a series of articles (e.g. 2001, 2004).<sup>43</sup> This now addresses the first problem with Locke’s account, the question of representational ontology. Slezak traces the problem back to Arnauld’s criticism of the analogy (as Malebranche supposed) between mental seeing and representations in the world that must have an observer for being seen. Slezak takes us to the edge of the problem with representation, which will be directly relevant to our critique of mind generally, and is why we should engage it. However, and this should be noted to begin with, he does say:

‘Despite its problems, the ubiquity of this view [that there is an intermediary representational state between subject and world] is undoubtedly a consequence of the fact that it seems impossible to imagine an alternative’ (2004; 164).

---

<sup>42</sup> Locke does treat of Clear and Obscure, Distinct and Confused Ideas in Book II, Chapter XXIX. But this is not the same as misrepresentation.

<sup>43</sup> These articles offer a background to the discussion, and reference to representative participants and their work.

This ‘impossible to imagine an alternative’ is what we must overcome for the alternative model we shall propose. But let us look positively for the moment on Slezak’s commentary. He quotes Fodor as saying:

‘It is, to repeat, puzzling how thought could mediate between behavior and the world.... The trouble isn’t – anyhow it isn’t solely – thinking that thoughts are somehow immaterial. It’s rather that thoughts need to be in more places than seems possible if they’re to do the job that they’re assigned to. They have to be, as it were, “out there” so that things in the world can interact with them, but also they have to be, as it were, “in here” so they can proximally cause behaviour.... It’s hard to see how anything could be both’ (Fodor: 1994; 83).

The conflict is between the intentional character of thoughts, in that they are about objects or states in the world and must therefore refer to them *accurately*, or truthfully (in Locke’s perceptual case, *by resemblance*), and yet at the same time exist in their causal role in behaviour *via* their representational instantiation. What kind of (impossible) ‘thing’ must thoughts be to achieve this? Slezak comments:

‘Fodor is firmly committed to a conception of mental processes as typically truth preserving. It seems we can’t do psychology with semantic notions [i.e. as internal causal states because they must accurately refer to the world (intentionally) *and must not misrepresent*] and we can’t do psychology without them [for they are the representational vehicles of mental causality, *psyche-ology* as already discussed]’ (2004; 170).

Since Fodor is committed to the causal function of mental processes, he finds he must discount intentionality. He concludes his 1980 paper:

‘My point, then, is *of course* not that solipsism is true; it’s just that truth, reference and the rest of the semantic notions aren’t psychological categories [as can be employed in a computational psychology processing symbols]’ (71).

We note that Fodor himself provides no ontological account of what thoughts are. Like Locke (with thought, but also perception), he assumes we know what they are from our own experience (or from the language of the topic), and can trust that when he is talking of them, they are an understood given.

To evade the difficulty Fodor raises, Slezak focuses upon representations as ‘used’ by a subject: they *seem* to be intelligible in their truthful character. However, when it comes to the *misrepresentation* of mental states, confusion arises because the truthfulness or otherwise of the representation is only relevant to the theoretician observing the process, not to the mind (or subject) ‘using’ them. The (theoretician) observer plays God (notionally) as viewing arbiter of whether the representation is true; but this is not the case for the ‘using’ mind (or subject). Slezak proposes that if the representation is seen to go wrong (to the God-positioned observer), it is a fault of the world, not the representation itself.

This is a highly compressed rendering of Slezak’s critique: he generalises it to various kinds of supposed problems in philosophy of mind. The confusion arises because theoreticians confuse the mental state of (e.g.) seeing with the corporeal representation of the picture (or *camera obscura*) that requires a seer. The mental state of seeing (*à la* Arnauld) is not concerned with establishing a true representation *to be seen* (*à la* Malebranche), but operates *within* its seeing framework. A seeing may be wrong (as judged by the external observer), but that is not the fault of the representation. I.e. that we see a lamppost and not a tree does not have to be regarded as a failure of *covariance* because the failure exists only in the *external* observer’s judgement, and is irrelevant to the system, i.e. we as viewer of tree/lamppost. In Slezak’s words:

‘The very problem of misrepresentation and, consequently, the entire constellation of semantic notions, are irrelevant to the enterprise of cognitive science.... Psychology has no obligation to explain why the world has gone wrong. Conversely, however, it has no obligation to explain why it usually goes right either. Hand-waving in the direction of evolution suffices here’ (2004; 173).

‘Hand-waving towards evolution’ means that seeing is a process which, in Ruth Millikan’s terms, fulfils its *proper function* as an adaptational operation; but that does not necessarily mean it will always get things right.<sup>44</sup>

We note that we (as seer) *might* reflect subsequently that we did mis-see the tree as a lamppost, for we now see the tree *per se*. But in so reflecting, we now assume the external observer position upon our own viewing.<sup>45</sup>

---

<sup>44</sup> For example, ‘Having a proper function is a matter of having been “designed to” or of being “supposed to” (impersonal) perform a certain function’ (1984; 17). It should be mentioned that Millikan’s bio-semantic view, as with Slezak implicitly here, is critical of consciousness and its facility of direct knowing. She refers to this as ‘meaning rationalism’. (See e.g. 91-93).

I will mention only one other of Slezak's generalisations. This concerns the industry around Hilary Putnam's Twin Earth thought experiment. Putnam's idea is that a person on earth has a belief about water as H<sub>2</sub>O when using the term 'water'. Transported to another world, identical to ours including the neurophysiological states of its inhabitants, our earth person retains his belief that water is H<sub>2</sub>O; but in this other world, the colourless odourless substance they call water is actually XYZ, the one exception to their identical nature. So on the other world, the inhabitants referring to water are actually referring to something different, XYZ. Putnam concludes from this that 'beliefs ain't in the head', but depend significantly on the external environment (1975). Thus was born the debate between externalism (*à la* Putnam) and internalism. Fodor is partial to internalism for causal reasons, but has flirted with versions of externalism.<sup>46</sup>

Slezak's view is that this is another example of the error involved in supposed misrepresentation. He says:

'Clearly, however, the [person's] mistake [in using the term water] may be unknown or unknowable to anyone except to ourselves who tell the story [i.e. us as external observer or theoretician].... My variant...may be portrayed equivalently by imagining that on this Actual Earth, God secretly switches all H<sub>2</sub>O to XYZ.... It is the world not the organism that is in error. It is only by illegitimately adopting the external omniscience of the theorist that we can formulate these puzzles' (2004; 176).

Slezak's critique is persuasive, particularly as it appears to be a way of dissolving many of the alleged problems of philosophy of mind. All one has to do is give up truth, not representation, in the notion of knowledge. We can say the mind has knowledge, or is knowledge, *via* its representational capacity, but we must not make truth the criterion of knowledge (as Schelling specifies) because correspondence is a criterion that presupposes a particular notion of mind – i.e. as intelligible in its constituents to the external theoretician.

'Asking about truth and its converse invites us to think of mental objects as external objects of apprehension and intelligibility themselves [as in Locke's notion of *resemblance* between idea and object]. This is of course the notorious error which has plagued theorising about the mind throughout its history.... Asking about the truth or

---

<sup>45</sup> This is an example that Slezak employs, particularly in his critique of Donald Davidson's position, see below.

<sup>46</sup> See his online review (2009) of Michael Tye's *Consciousness Revisited* (2009) in the *Times Literary Supplement*.

falsity of a representation is inescapably to invite understanding of its meaning.... It should be clear that the role of mental symbols and our scientific concern to explain this must be independent of whether we are able to understand the symbols as meaningful to us' (2004; 179).

Locke's notion, without his realising, attempts to impose understandability on the symbol (the primary quality perceptual representation) by its resemblance to the object.

### 3.1.5 Slezak's unintended destruction of consciousness

But as said initially, Slezak's critique only takes us to the edge of the problem, persuasive though it is. It does so for two reasons.

Firstly, to deny that *truth* is a criterion of consciousness for the purposes of saving representation is to deny that consciousness exists based upon saving the mission statement of cognitive science. Slezak undermines consciousness (or mind) by rendering it so *non-divine* as not to be knowledge at all. This might be a valid step on the way to locating a biological function for the state *misunderstood* as consciousness. But Slezak does not entertain that possibility.

It might be said that the Slezak/Arnauld position enables us to escape the regressing homunculus problem. We seem to know. But how do we know? Nothing available to our reflection explains how we know; it's a mystery. Thus *knowledge* seems to require a knower of the material (as it appears in our head) that we know. This is the ubiquitous homunculus who *effects* our knowing. But what effects the knowledge of the homunculus? Another homunculus. Etc.

A direct knowing (*à la* Arnauld) might evade *this* problem, but it does not explain *itself*: it does not explain what conscious knowing *is*. It is not that we need a knower of what is known (the homunculus): we need to understand what knowing is *ontologically*. E.g. *how* does Slezak's supposed mental subject or external observer observe exactly – even if not the truth? He does not tell us, any more than Locke tells us what inner representational perception is.

Indeed, Slezak still proposes that: ‘Undoubtedly the close coordination of internal representations with the external world is required for intelligence as an adaptive feature of organisms’ (2004; 176).

But why? This simply keeps in the brain what has been supposed as mind, but denies we can understand it. If we cannot understand its contents (mental symbols), why should we continue to suppose it *is* mentality (e.g. causal representation to a subject)? For the ‘close coordination’ need not be associated with mind and its intelligence, which Slezak unquestioningly supposes.

But secondly, Slezak’s approach connotes another blow for consciousness, though again he appears not to register it. He applies his critique to Davidson’s contention ‘that a creature cannot have a belief if it does not have a *concept* of belief’ (2004; 177). He quotes Davidson:

‘Someone cannot have a belief unless he understands the possibility of being mistaken, and this requires grasping the contrast between truth and error – true belief and false belief’ (Davidson: 1975; 22).

What Davidson is actually proposing, functionally, is that we *must* be able to know the contents of our own minds by reflection. If unable, how could we register the belief we have as true or false? But Slezak says: ‘It seems very unsatisfactory...to deny that animals might have beliefs because they are unable to know that they have them and reflect on their truth value’ (2004; 177).

There are two things at issue here. The first is that Slezak commits the very error he attributes to ‘theorising about the mind throughout its history’. For it can only be said of animals that they have beliefs *by the external observer*. If animals cannot know they have beliefs, they do not have them. For beliefs are (1) mental states which must be available to reflection, and (2) must be expressible (or thinkable) in language, neither of which most animals have. We can no more claim the cat believes there is a mouse in the hole than that the cat can tell us so *precisely* because of Slezak’s own analysis. Of course *we* can share the discourse that the cat believes there is a mouse there (as external observers), but this is irrelevant to the cat, so our discourse is valueless as *science* of the cat or the cat’s brain concerning beliefs.

But the more significant issue is this. Slezak's analysis, insofar as it is valuable, undermines the notion of reflection – and thence consciousness. Here is Arnauld's declared position which Slezak (presumably) supports.

'Our thought or perception is essentially reflective upon itself.... For I do not think without knowing that I think. [I.e. I know I am conscious because *my* consciousness tells me so.].... As well as this implicit reflection, which accompanies all our perceptions, there is also something explicit, which occurs when a geometer finds in examining his perception of a triangle that, having conceived a figure bounded by three straight lines, it must have three angles...' (1990; 71).

Though Arnauld claims 'there is nothing in these two points that can reasonably be contested' (*ibid.*) (a confidence at odds with the issues, cf. Locke), 'reflection' *demands* explanation.

Firstly, implicit reflection is not merely a taken-for-granted *condition* of consciousness: it is its *justification* (cf. Kant's apperception). As with Locke, we are presented with a *fait accompli*: the entity (or state) the theorists propose, i.e. consciousness, self-justifies in our experience. Far from being uncontestable, this is extremely suspicious, for it countenances no alternative. As said, we shall address this issue in the next chapter.

Secondly, however, concerning explicit reflection: How can the geometer examine his perception of a triangle? What does *examine* mean? Does it mean 'look at his perception and interpret it'? But Arnauld (and Slezak, therefore) tell us that the subject does *not* look *at* representations (these 'chimeras'). Perception is an objective *act* (or modification) of the soul. Arnauld's account seems to require making the three-sided percept itself an external object for perceptual examination of *it* (the very thing he has dismissed in Malebranche), thereby to find three angles.<sup>47</sup> This is, of course, impossible. So unless we succumb to the unexplained internal mental interrogation of perceived objects by Arnauld, we are without an account of explicit reflection in direct conscious seeing, thus rendering direct consciousness, as it were, inert.

Since we seem to have run out of convincing alternatives on consciousness models, we might now descend to the realm of neurobiology and its *production* of consciousness to

---

<sup>47</sup> The reader will be reminded of the preceding discussion concerning the (un-)availability of *qualia* as objects of knowledge.

see if the claims made for consciousness can become clear from the bottom up, as it were. But now we encounter a different kind of problem. In his criticism of Ned Block, Slezak makes this point:

‘Block distinguishes between the “autonomous” and “inherited” meaning [of representation] saying “The representations on the page must be read or heard to be understood, but not so for the representations in the brain.” (Block: 1986; 83) ... However, despite this explicit recognition of the distinction, the discussion is conducted in terms which tend to obscure or blur the crucial differences. Thus apparently talking about the meanings of inner representations, Block asks “what is it to grasp meanings?” (82) and “what is it for the brain to grasp meanings?” (83). The idea of a grasping or understanding here is exactly the wrong one in relation to internal representations which are precisely not grasped or understood’ (2004; 183).

Slezak makes the distinction between the mind grasping the meaningful (the world perceived, *à la* Arnauld) in its luminous intelligence, and the ‘dark’ electro-chemical brain operating without a grasp of anything (for it is merely electro-chemistry). But the evident question then is: How and why does the brain *create* ‘readings’ and ‘hearings’ as understanding/meaning (us as consciousness) which cannot grasp the processes of their own production? How, in Nietzsche’s words, does the rational arise from the irrational (1986; 12)?<sup>48</sup> Surely we are now at the ‘extreme’ position of Colin McGinn who has said that:

‘We need to cultivate a vision of reality (a metaphysics) that makes it truly independent of our given cognitive powers as a proper part.... A deep fact about our own nature as...embodied consciousness is thus necessarily hidden from us’ (1995; 289-290).

I.e. we are *incapable* of understanding consciousness.

McGinn’s position dismays (probably) the vast majority of philosophers and cognitive scientists. Nothing, it is hoped, should be beyond our cognitive reach, and certainly not cognition itself. What, then, would be the point of cognitive science? Yet this is the impasse we appear to have reached.

---

<sup>48</sup> In some conformity with this issue, Kant says in the chapter of the first *Critique* concerning Schematism: ‘This schematism of our understanding, in its application to appearances and their mere form, is an art concealed in the depths of the human soul, whose real modes of activity nature is hardly likely ever to allow us to discover, and to have open to our gaze’ (1933; 183).

For if the very *idea* of knowledge depends upon the fact that we know we have it (Arnauld, Locke, Block, etc.), and the brain disallows our knowing how this is accomplished – as Slezak and McGinn say in different ways, and Block in one apparent breath –, then clearly the *ground* of knowledge is beyond us. We cannot be the understanding observer of our own *neural* processes causing ‘us’.

But now the question must arise: If we cannot know *how* we know anything, how do we know we *do* know anything? How do we know we are *conscious* if the knowledge of being conscious (as reflection) is unknowable, as is consciousness itself? We cannot simply assume knowledge if we cannot demonstrate it in ourselves or indeed define it. This is where Slezak’s critique actually leads us. (We shall return to this in the next chapter.)

## 3.2 Illumination

### 3.2.1 The illuminating topic of colour

Let us move to the interface between the theoretical and the empirical by engaging the topic of colour. If the universe is only physical, then it makes no sense to talk of states of the world as mental as distinguishable from the physical. It certainly makes no sense to refer to things as mental on the supposition we are being meaningful regardless of the fact that we have no means (in any agreed way as Kuhn’s ‘normal science’) of specifying how the mental can be physical. Yet this we find throughout the literature. Concerning colour, according to Peter Ross: ‘The problem can be put as follows: What sorts of properties are the colors? Are they *physical* properties of physical objects, *mental* properties, processes or events of perceivers’ visual states, or perhaps some other sort of property?’ (2001; 42).<sup>49</sup>

If we see colour (which we appear to), and it does not belong in the world (as Descartes thought), where is colour, given that our brain, which must be the source of the colour we see, is an object in the world and cannot *contain* colour? (As an example of how wrongly this can be construed, Frith states: ‘Color is in the brain, not in the world’ (2007; 134). Thereby the brain is not in the world, which is an unfortunate turn of events.)

---

<sup>49</sup> Gary Hatfield adds a third option. ‘Relationists...argue that color, considered as a property of objects, is a relational property; it is a property that surfaces and light sources have of causing experiences with various phenomenal characters in perceivers’ (2009; 281). From our point of view this does not add to the investigation.

According to Ross, the modern debate has been considerably inspired by two writers: C.L. Hardin (1993) whose position has been that of *subjectivism*, or that colours belong in the mind, and David Hilbert (1992), who claims that colours are the physical properties of physical objects, *physicalism*, namely the reflectance properties of the surfaces of physical objects.

The two sides of this debate present a *dichotomy*. It would appear we must make a choice between two positions, one that takes account of the physical universe, the other that emphasises the mind. This dichotomy, it will be argued, is false. For (overturning Slezak's position that it is 'impossible to imagine an alternative'), it is precisely the alternative that we will identify. However, it is instructive to consider what is involved in the dichotomy, and why we need a third way.

Ross attributes the following classification to Austen Clark (1996, 2000). It will be now quite familiar.

In one case the visual field is regarded as an array of mental impressions. I.e. the visual field is a mental medium in which objects and properties are represented. This is commonly termed *indirect realism about perception*. This, of course, fits the Malebranche approach.

In the other case the visual field is an array of physical objects and properties as seen. What is represented are not mental impressions, although there is mental processing to arrive at what is seen. This is commonly termed *direct realism about perception*. This is the Arnauld approach.

It is hoped the reader will be alerted to the potential flaws of these approaches. Both assume that colour *per se* is apparent to the perceiver, whether the perceiver views mental impressions, or the physical world as seen *via* mental processing. But the perceiver, or the perceiving, are unknown kinds of (knowledge) category, and depend upon the (mental) reflective means of justification (*à la* Locke).

Ross says, concerning 'perceived' objects: 'It seems that we must individuate these particular parts [mental representations, or objects as seen] in what I have called the natural way – in terms of the particular sensed locations, and specifically in terms of relations between perceivers and particular sensed locations which are particular parts or wholes of sensed objects (or regions of physical space)' (2001; 52).

The issue here (the so-called *binding problem*), as Ross presents it, is that *processing* in the brain of the elements of what is experienced occurs in stages and at different locations, whereas what we see is organised and structured as a unified field comprising individual parts as they would be in the real world – as perceived by a perceiver (notionally). Any successful account must address this.

To repeat, however, the ‘fact’ of the ‘perceiver’ and the ‘fact’ of the ‘perceived’ are what we might term both metaphysical and scientific assumptions, and this consideration should be retained throughout the following. (The initial ‘It seems’ of Ross’s statement itself requires validation, which he does not offer.)

The problem, according to colour science,<sup>50</sup> is that what we *see* as a determinate colour cannot be directly correlated with an external surface, for indefinitely many different surfaces can generate the (apparently) same colour for the perceiver. Moreover, no physical properties exist to which the classification of colours as qualitatively similar or dissimilar corresponds. Thus there is a serious disjunction (and discontinuity) between what we see from objects, and what is there.<sup>51</sup> The once assumed continuity is overturned by science. *Prima facie* this seems to support the view that colours are a characteristic of the mind, i.e. the way we see the world *mentally* as opposed to the way the world is (cf. Locke’s secondary qualities).

### 3.2.1.1 The evolutionary factor

But now we need to introduce another factor, one already alluded to in the Slezak discussion. This concerns what our experience is *for*, in evolutionary terms. If we take consciousness as ‘giving us’ what is in the world as knowledge, then *de facto* it should give us what *is*, what is true. Where this is not the case, as with colours, we may be drawn to the oversimplified notion that what we experience is *subjective*, or conversely is *not objective*. On the other hand it may be that what we experience results from the *role* the experience performs for us as biological organisms: i.e. ‘truth’ is only relevant relative to our biological (evolutionarily determined) existence. Surely, theoretically, we might suppose that ‘genuine’ knowledge

---

<sup>50</sup> ‘Discussions about color... in ignorance of the visual science are intellectually irresponsible’ (Hardin: 1993; xvi).

<sup>51</sup> The seeing of reflected light from objects is different from seeing a direct light source; it is influenced, for example, by adjacent surfaces.

would be an evolutionary benefit but, apart from the fact that we would have to be able to conceive what such might be (philosophers have tended to equate knowledge to the ‘genuine’), we must accommodate the fact that we are biological organisms, and how we are physically created serves the function of evolution, which *per se* has no ‘interest’ in, nor plausible means for generation of genuine (i.e. true) knowledge. (As Markus Peschl states it: ‘The goal...is not to reconstruct the environmental structure as accurately as possible, but to provide the organism with relevant information and representations for generating adequate behaviour’ (1999; 208).)

But there is more to this. Colour science begins, as it were, with the eyes of organisms. For humans and close primates, colour discrimination is trichromatic. This results from the three colour receptors, or cones of the retina, picking up short, medium and long wavelength light (approximately). Other species have a different vision: more complex are many birds with tetrachromacy (four cone types), while pigeons are conjectured to be pentachromats. Mammals including other primates are generally dichromatic, and marine animals monochromats (cf. Kelber, *et al.*: 2003). This variation, it is hypothesised, is evolutionarily determined for survival value in the environments to which the species are adapted.

However, are we to suppose marine animals are *conscious*? That as organisms or brains, they sustain subjects for whom *colours* are represented? Certainly we may suppose their brains process signals from their visual apparatus and make appropriate behavioural moves. But we cannot (surely) simply wrap into colour science the notion of human consciousness. Yet Ross states we must ‘individuate’ objects in a ‘relation between perceivers and sensed locations’. An evolutionary continuity for vision between marine animals and humans (surely), however, cannot be a conscious one requiring colour *perceivers*.

Indeed, the most complex colour vision on the planet belongs to the mantis shrimp, or *stomatopod*, which has up to twelve different spectral receptor types (Cronin & Marshall: 1989; Caldwell: website). The enormous power of the shrimp’s eyes, it is conjectured, is for detecting prey and avoiding predators. But can we possibly suppose there is a conscious

subject in the shrimp's brain (cf. the size of the human brain and primary visual cortex) for whom colours (and the other vast array of visual information) are *represented*?<sup>52</sup>

Mohan Matthen, commenting on Ross's article, states that:

'The colors we detect are not necessarily the target of avian or apian color vision. The physicalist can, and should, take this variability on board. Physical properties are numerous, and it is possible that different color perceivers fasten onto different properties or even property structures.... Honeybees use polarized light color to identify flowers, and birds use it in directional navigation. Such uses of color are unavailable to us. The color categories generated by such activities are, again, species-specific' (2001; 123).

But this emphasises the question without identifying it: Are our uses of the words 'color' or 'perceiver' meaningful? For colour, like light, is a *mentalist* notion. What impacts the eyes of organisms is electromagnetic radiation, not light (or colour). What is transported to the brain from the eye, or even remains somewhat within the functioning of the eye (in some species), are physical signals which have nothing to do (surely) with perceived light or colour. Yet between shrimp and humans something appears to be introduced involving colours and perceivers, *viz.* consciousness. But does it?

And there is another factor to consider: *colour constancy*. Objects seem to retain the colour associated with them in 'normal' conditions, even if the conditions are not normal (e.g. between noon white light and sunset red light predominance).<sup>53</sup> This seems to imply an eye + brain function *superadded* to mere colour but unrecognised (even reflectively!) by humans in their normal vision (i.e. only under experimental conditions). What seems evolutionarily important is the (biological) *use* of colour in the process of identifying objects, rather than colour as such (i.e. it is *not* superadded but functionally intrinsic). Therefore, even with humans who do *see* colour (supposedly), have we even the right conceptual notion of colour as a biological category? Are we, in proposing *colour* exists (differentiated from its use), presupposing an access to the nature of our subjective states that is illusory because the biological role of colour is *not* self-evident to us in our seeing? Moreover the constancy

---

<sup>52</sup> This problem is outlined, but not solved, by Jonah Lehrer, (2011; 62-67)

<sup>53</sup> This is according to the *retinex* theory (*retina* + *cortex*), as formulated by Edwin H. Land (e.g. 1977). However, the phenomenon was already known, and Land was criticised for not referencing it.

condition could exist with species that do not *see* colour at all, and they certainly have no reflective capacity concerning their ‘experiences’.

It is noteworthy that Matthen states:

‘However you interpret the ontological significance of “color constancy” – whether you take it as evidence that color vision is after the truth or take it, on the contrary, as evidence of its mendacity – there will be undeniable correlations between *perceived* color states and environmental situations involving reflectance, luminance, figure and motion’ (2001; 123).

Given his previous statement about the evolutionary significance of colour, and its species-specificity, this statement is evidence of an ontological confusion. The notion of *truth*, as we saw from the Slezak discussion, is highly suspect. Certainly there is no external observer *of* truth. And the statement assumes, in terms of its ‘undeniable correlation’ between ‘*perceived* color states’ and ‘environmental situations’, that colour is perceived – at least by humans. This ‘undeniability’ has the same exaggerated ring of confidence as Arnauld and Locke, for it can only result from the notion of (mental) reflection in which not only (true) knowledge, but knowledge of knowledge as known avails.

Does the apparent existence of colour constitute knowledge to a perceiver (who has not as yet been explained), or mislead into a false conception of perception, or whether indeed perception exists at all?

### **3.2.1.2 Rey, Ross, and the unacknowledged end of mentalism**

We are now in a labyrinth from which there has been no escape. Georges Rey says:

‘Although we are sometimes subtle and insightful, a moment’s reflection reveals we know next to nothing about even such basic activities as perception, thinking, reasoning, language, decision making, motor control, not to mention consciousness, creativity, scientific insight, or morally responsible action’ (2007b; 74).

One might imagine from this he would reject mentality *tout à fait*. Find another way. But no. He theorises as follows concerning intentional objects (i.e. ‘objects’ of the mind):

‘I don’t mean that they must be reckoned in one’s ultimate ontology. While they are “objects” of our thoughts...purely intentional objects don’t exist, nowhere, nohow’ (2007a; 116f).

If the ‘objects’ of mentality do not exist, surely *indeed* it implies the end of mentalism. But at the same time he says: ‘My proposal...is to treat qualitative experiences [of e.g. colour] as involving causal-computational relations to restrictive representations, the specialness of the representations consisting in, *inter alia*, their serving as the output of sensory modules’ (115).

But what kind of beings have outputs of sensory modules *experienced* as colour, or indeed intentional ‘objects’ as seen or thought? Rey even says elsewhere ‘each person would have available to them representations’ (117). Where and what are these *persons* supposed to be, and how do they ‘have available’ representations?

Having made questionable anything in mentalist discourse, Rey reverts to structures entirely within that discourse, e.g. experience of colour, or representations to persons. Indeed, what he refers to is not the *discourse*; he refers to the *knowledge* that we do not have ‘upon a moment’s reflection’. But if we did have knowledge, of what would it be? What ontological status would it obtain? For Rey’s position seems to rule out knowledge entirely; its content would ‘not exist, nowhere, nohow’.

Ross, on the other hand, is determined to eliminate mentalist notions. Having dismissed the subjectivism of colour, he concludes:

‘We can *distinguish* between...neural processes and the sensed colors themselves.... We can allow, for example, that the same neural process gives us perceptual access to indefinitely many metamers. And although these physically distinct properties [metamers] are experienced as the same sensed color, we are not required to conclude that sensed colors are themselves neural processes. Rather, we can hold that sensed colors are disjunctive physical properties of objects to which we have perceptual access in virtue of certain neural processes’ (2001; 55-56).

Rey has virtually written off mentality, though without denying it exists. Ross is more absolute. For him there is a transparency of the world to us, in particular colours. ‘[We have access to] sensed colors...in virtue of certain neural processes.’ How is it achieved? Neural

processes just ‘give’ it. Philosophical assumptions, he claims, have led us astray concerning colour subjectivism.

‘Considerations about color perception bring us back to what I think is the commonsense view that colors really are located on physical objects we perceive as well as open the possibility that colors are describable in physical terms’ (2001; 56).

But what Ross does not tell us (in his commonsense), is exactly how neural processes, actual physical states, *could* ‘give us’ the colours that exist on the object; nor indeed what ‘us’ this is (which neural processes also must create) who perceives them. If Ross is endorsing an Arnauld-type direct realism, effectively proposing that mentality as an intervening modus is transparent, he must yet explain how the ‘colour as seen’ *derives* from neural processes (themselves uncharacterised). For not only is this a mighty (or superhuman) task, he appears not to appreciate that, without *explaining* the derivation, colour remains a purely mentalist notion and cannot *by definition* be on objects. Having rejected mentalism, Ross must explain what he does not explain. Rey, on the other hand, puts it in an inverse way: ‘How come no one else can enjoy my reddish experiences by inspecting my brain?’ (2007a; 117).

Neither Rey nor Ross consider whether seeing colours is actually a continuity from marine creatures to humans. Why, for example, would purely neural processes (we assume) in the mantis shrimp (for Ross) become *perceived* colours in the human being? On the other hand, what (for Rey) is the change that makes neural processes (causal-computational, if you like) different for mantis shrimps and humans *given that* humans appear to perceive colour? Not only does neither account answer these essential questions, they are not even considered.

Between Rey and Ross, it is the former who holds a healthy scepticism about his own position. His position is ‘a “strategy”, not anything like an adequate theory, since...there is...a great deal...that needs to be spelled out’ (2007a; 112f).

He states further that his causal-computational theory of thought is ‘not so much a serious “theory,”...but...a promising research program, within which, however, quite independently of qualitative experience, there are plenty of explanatory gaps’ (115). What might surprise us, by now, is that Rey feels the need to save mentality at all and not look for a more effective way of approaching the brain from its *physical*, or neurobiological origins.

Why not search for a theory without explanatory chasms which fulfils some genuine biological function?

### 3.2.2 Summary

Dennett says: ‘Color has always been the philosophers’ favourite example’ (1991; 371). This is because the problems it presents characterise (1) the notion of mind, and (2) its intractability to explanation.

But colour is merely symptomatic. For example, Rey says in a footnote: ‘My favourite examples are jokes: how could one hope to explain regularities in people laughing at jokes without adverting sometimes to their content?’ (2007a; 118f).

A scientific explanation of human beings as purely physical organisms need not explain colour *per se*, i.e. as a real property. Nor need it revert to the content of jokes to explain the regularities of people laughing at them. Scientific history is filled with examples that entail ‘having to revert to’ some feature, which resulted from working on the wrong model. For example, life invoking vitalism or (Henri Bergson’s) *élan vital*.<sup>54</sup> But while the science or philosophy of mind, as opposed to brain, accepts this supposition, there will be no progress. It will continue to defeat scientific analysis.

### 3.3 Empirical findings

The first section in this chapter was theoretical. The second engaged both theory and empirical findings. In this third section we shall focus on the empirical.

Of course the literature of the empirical, as with the theoretical, is enormous. We cannot treat of all findings (as we cannot treat of all theory). Our approach will be to undermine the supposition of consciousness as evident from some well-known examples. The reason it has not been so understood before is that there has been no alternative theory (as Slezak has suggested).

---

<sup>54</sup> Though Bergson (1911) actually criticised the notion of vitalism, his *élan vital* has often been associated with it.

### 3.3.1 The consciousness structure

The structure of consciousness as a causal property is as follows:

A:                    input            consciousness            action

Consciousness is supposed to play a causal role between input to the organism and its action. Consciousness *mediates* action.

This model states that I cannot act unless I see what it is that causes my action. On the other hand I cannot act unless I have thought the thought that will cause my action – or felt the emotion, and so on.

Of course the model does presuppose prior experience. I get out of the way of a car moving towards me; but I have prior knowledge of its likely impact upon me.

It became evident that action did not necessarily involve conscious influence. This impediment to the model resulted in the formal adoption of the notion of the unconscious or non-conscious, as already discussed. Thus we have:

B1:                    input            consciousness            action

or

B2:                    input            unconscious            action

There was some furore at the notion that actions could be caused without the subject being aware of their cause, particularly when associated with Freud's theories as to why they were unconscious, *viz.* the subject could not acknowledge the causal reasons. Yet one might have supposed, as already suggested, that a firmer stand could have been taken then as to whether consciousness was ever a causal property, or existed at all.<sup>55</sup> Epiphenomenalism, the notion that consciousness is not causal, is not, of course, the same as the notion that consciousness does not exist.

But now the unconscious had gained respectability, all sorts of dubiousness under the guise of explanation could be conducted. The first was already extant: *habit*.

---

<sup>55</sup> This is a significant topic in its own right. Freud was influenced by Georg Groddeck, who he mentions in *The Ego and the Id* and the *Introductory Lectures*. Groddeck wrote a book titled *Das Buch vom Es* (1923), or *The Book of the It* (1961). As Freud says in the former: 'I am speaking of Georg Groddeck, who is never tired of insisting that what we call our ego behaves essentially passively in life, and that, as he expresses it, we are "lived" by unknown and uncontrollable forces' (1962a; 13).

Locke says:

‘Action is the great business of mankind, and the whole matter about which laws are conversant’ (1997; 266). ‘[The] power or ability in man, of doing anything, when it has been acquired by frequent doing the same thing, is that idea we name *habit*’ (*ibid.*). Indeed ‘*habits*, especially such as are begun very early, come, at last, to *produce actions in us, which often escape our observation*’ (145).

This account has survived into the modern day. ‘Acquisition of a complex skill can make enormous demands on consciousness. As the skill becomes practiced, however, attention and conscious control gradually absent themselves’ (Allen and Reber: 1998; 317).

The authors do not tell us what consciousness has done *that* we acquire a complex skill (we must assume that from our own experience); nor do they tell us (consequently) what ‘attention and conscious control gradually absent themselves’ means in terms of the unconscious taking over. I.e. why *exactly* do habitual skills, once acquired, not require conscious control? This account is simply convention.

Perhaps (as we shall come to) there is an entirely different explanation from consciousness/unconsciousness, which has survived despite an absence of credible detail. For if the brain programmes itself to acquire a skill, why does there need to be an intervention of a (functionally/ontologically unexplained) consciousness, as in A. Indeed, if there is a possibility of B2, as is apparently well-established, why have A at all?

But there is more to say at this initial stage. It is apparent from A that what actually is proposed is that physical input is converted *to* consciousness to become causally effective *as* consciousness, and then whatever consciousness causes must be effectively communicated to output ‘modules’ to enable appropriate action. Now quite apart from the fact that we have no account of what neural states are as conscious states, we must question what benefit could be gained by the conversion of physical neural states to (physical) conscious states. (Dennett (1991), of course, proposes that conscious states are the accumulated result of purely neural states.)

If we suppose our experience as consciousness is of a different ontological order from the ‘merely’ physical (mass/energy space-time, electro-chemical) which is not consciousness, how could there be effective *communication* of physical input into consciousness and from consciousness back to physical states to generate action? For example, suppose in seeing a

tree I want to go and touch it. What physical communication method is there from a seeing and a wanting to the motor functions of action? The translation seems no less difficult to conceive now than when Descartes proposed his original semantic translation model from the physical to mind and back.

Of course, one could write this off (as Dennett no doubt does) and claim that it just happens. Bernard Baars, for example, is dismissive of the difficulty in his Global Workspace Theory of Consciousness. He states that:

‘GW (Global Workspace) theory gives the most complete account to date of the interplay of conscious and unconscious processes in perception, imagery, action control, learning, attention, problem-solving and language. These topics can all be usefully treated as types of information processing’ (1996a; 211).<sup>56</sup>

Unfortunately he gives no account of *how* there can be an ‘interplay’; support for his theory derives entirely from our own ‘experience’, to which he constantly appeals (as Locke, *et al.*).

But we do *not* experience unconscious processes. And he does not define consciousness. It is, for example, simply tautological to say that perception is (a facet of) consciousness, if perception is not defined in any other terms than consciousness. Dennett correctly says: ‘If your model of how pain is a product of brain activity still has a box in it labelled “pain,” you haven’t yet begun to explain what pain is’ (1991; 454-455) – pain *per se* depending upon being conscious.

Baars’ model, and its inherent difficulty, is but one version. Here is Block on *access-consciousness*, associating it with Stephen Stich’s *inferential promiscuity*.

‘A state is access-conscious (A-conscious) if, in virtue of one’s having the state, a representation of its content is (1) inferentially promiscuous (Stich, 1978), i.e. poised to be used as a premise in reasoning, and (2) poised for [rational] control of action and (3) poised for rational control of speech. (I will speak of both states and their contents as A-conscious.)’ (1995; 231).

---

<sup>56</sup> See also 1996b.

Block's distinction of access consciousness from phenomenal consciousness is an arbitrary expediency. It allows him to suppose (unjustifiably) that the work of consciousness happens without the inconvenience of an unexplained phenomenality.<sup>57</sup>

Since neither Stich nor Block have any idea (for no evidence is forthcoming) how a conscious state could be poised as a premise for reasoning, rational control of action or speech in any describable physical/functional sense, we take it that this is philosophers' discourse on the benefits to their theoretical positions if such were so. But this is just Plato's illumination myth clothed in pseudo-scientific formulation.

### 3.3.2 The breakdown of the brain phenomenon

Let us move to more direct matters. It has become evident, since the famous incident of Phineas Gage, that a personality change could result from brain damage (Harlow: 1848). The incident occurred because the tamping iron he used in his work on the railroad was driven up through his cheek and out from his skull by an explosion, removing part of his prefrontal cortex. From being a gentle, reasonable and caring person, he became brusque, selfish, and subject to fits of profanity.

Gage suffered brain deficit, and the many forms of deficit have become a major source of information concerning the relation of parts of the brain to features of (so-called) consciousness. Indeed, the notion of the unity of consciousness, or perhaps a better expression would be the integrity of consciousness, has been undermined. Whatever consciousness may be supposed to be, it is an *assemblage* of many different facets, and the absence of brain function in areas can manifest in the absence of facets of that neural phenomenon.

The question for us, however, is not one of pure description: there is a vast literature on brain deficit. The question is: What is illustrated thereby *vis-à-vis* consciousness – or not?

This is of the utmost importance, since in a substantial section of the literature, deficit is deemed to indicate the negative impact on *consciousness* (and its facets) with no consideration as to whether consciousness is in fact relevant.

---

<sup>57</sup> In the referenced paper, Block terms consciousness a *mongrel* concept. He terms it elsewhere a *cluster* concept. Chalmers (1996) also refers to the many different uses of the word consciousness. What this indicates is that neither has any scientific theory of the brain phenomenon.

### 3.3.3 Milner and Goodale: Two visual streams of the brain

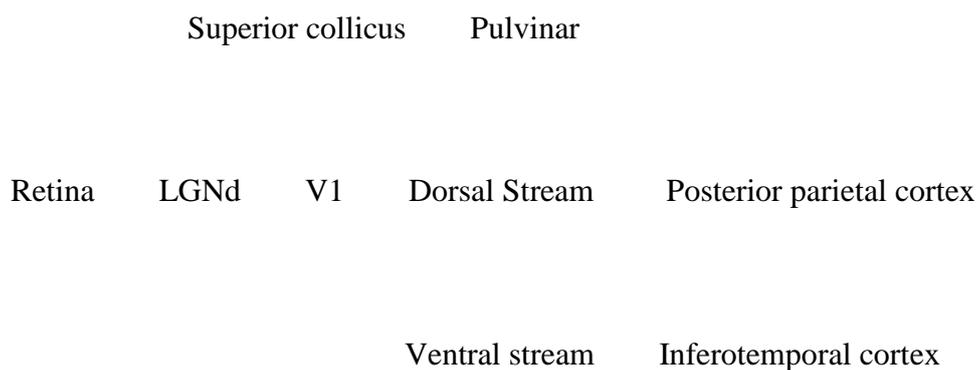
David Milner and Melvyn Goodale published their text *The Visual Brain in Action* in 1994. A second edition was published in 2006. They published an intervening text in 2004 titled *Sight Unseen*.

Their findings are regarded as a major contribution to the neuroscientific literature. Progressing the work of Mishkin and Ungerleider (1982) on two visual cortical pathways, Milner and Goodale investigated what are termed the *dorsal stream* and the *ventral stream* of visual input through the brain. These are colloquially referred to as the *where* and *what* pathways.

The two streams are associated with the functions they perform. The dorsal stream is for ‘unconscious control of action’, and the ventral stream for ‘conscious perception’ (2004; 31).

In other words, it appears there is not one function of consciousness mediating input and causing output as in A. One might say there is a B1 *and* a B2, except that B1 is not the cause of action, at least not directly.

Graphically these streams may be represented thus:



As described by Milner and Goodale:

‘The ventral stream receives most of its visual input from the primary visual cortex (V1), which in turn receives its input from the lateral geniculate nucleus (LGNd) of the thalamus. The dorsal stream also receives input from V1, but in addition gets a

substantial input from the superior colliculus *via* the pulvinar, another nucleus in the thalamus' (2004; 49).

The significance of these two streams is found in deficit. The subject with whom Milner and Goodale worked, who they refer to as Dee (or D.F.), suffered brain damage in the ventral stream as a result of carbon monoxide poisoning. It was apparent, however, that Dee could still perform complex manipulative skills which required the brain to have access both to the world as the target of those skills, and to motor control to effect those skills in test conditions. But Dee had no conscious awareness of the form or shape of objects in the world upon which she was operating, nor of the effectiveness of her motor actions. Dee suffered from what is termed *visual form agnosia*. The impairment in her particular case, fortunately, did not prevent her seeing colours or surface texture, and her hearing and touch were also effectively normal. As they say: 'The most amazing thing about Dee is that she is able to use visual properties of objects such as their orientation, size and shape, to guide a range of skilled actions – despite having no conscious awareness of those same visual properties' (2004; 29).

It might be said that all that had happened was that Dee had suffered damage in her (single) vision system. However, Milner and Goodale bring forward complementary examples to demonstrate this is not the case. These are where the subject has apparently normal visual consciousness, yet cannot effect action. They result from damage in the parietal lobe of subjects, i.e. in relation to the dorsal stream. The condition is termed *optic ataxia*.

Reviewing the substantive material they bring forward is most interesting. However, our focus must be upon the implications of their findings, and how they themselves interpret them. For at first glance, it might seem that the ventral stream is irrelevant to action (given Dee's capabilities). Even, that visual consciousness is epiphenomenal. For after all, as they say regarding evolution: 'There was never any selection pressure for "internal picture" shows – only for what vision could do in the service of external action' (2004; 40). The orientation of this comment is that of Locke, except, of course, Locke supposed there was one visual system, and it was consciousness (unless habit had taken over).

However, they do not propose the ventral stream, or consciousness, has no impact upon action; and they propose that consciousness serves other functions than action.

What is important to grasp at the outset, as we shall come to, is that Milner and Goodale assume that the ventral stream essentially serves consciousness, rather than that it may be understood to serve both as a neural function *per se*, and the generation of the neural phenomenon which has been supposed as consciousness. This is in spite of their comment about certain tests requiring communication skills at which Dee failed. ‘Dee could do none of these things – not because she couldn’t communicate, but because she had nothing visual to communicate. She had no conscious...visual experience...to share with us’ (2004; 29). This remark will be of the utmost importance.

Miner and Goodale’s posture is that consciousness provides us with a mental world. The fact that a mental world is ontologically of an apparently quite different order from neural stuff does not engage their attention. (They are psychologists/neuroscientists not philosophers – although it is evident that the approach of one *or* another is not enough.) Here is a crucial example of their interpretation of their findings, taken from their most recent text.

‘Not only do the two streams use different frames of reference but they also have different temporal characteristics. The dorsal stream may enable us to reach out and grasp objects with exquisite ease, but it appears to be trapped in the present.... The ventral stream, in contrast, allows us to escape the present, and bring to bear visual information from the past’ (2006; 245-246).

Thus concerning optic ataxia (consciousness without the ability of action):

‘Recent work has shown that such patients show a paradoxical improvement of reaching accuracy when action is delayed.... Patients with dorsal stream damage can still use visual information to guide their behaviour, by using the ventral stream as their route to action. We would further argue that this indirect route to action is not just a fall-back, for use when the system is damaged – but rather a good example of how the ventral stream contributes to the control of action by bringing the past into the present’ (2006; 248).

This ‘past’, according to Milner and Goodale, is the mental life created out of the ventral stream. But is it mental life actually? Here is their characterisation of vision upon mental life.

‘It is through seeing that we gain most of our knowledge about external reality, and the possession of that knowledge, in turn, powerfully affects the way we see other things. In fact visual knowledge determines much of the basic content of our consciousness.

Visual knowledge allows us to plan actions, to picture the consequences of those actions, and to relive (sometimes with pleasure, sometimes with regret) what we have seen and done in the past' etc. (2004; 39).

Here is the cultural supposition we are undermining. This is not to say that past programming from the ventral stream as neural states *per se* does not influence present actions. Of course it does, since it determines the majority of all we do *precisely* by pre-establishing it as 'learning' of the world. (In Edelman's felicitous phrase, a 'remembered present' (e.g. 2000).) But that does not mean it is mentality.

Here is Milner and Goodale's account of learning, which will be reminiscent of Locke and Allen and Reber.

'When an unfamiliar action is first performed, the movements are typically performed in a deliberate and cognitively controlled fashion. After the action has been performed a number of times...these slow cognitive control processes seem to get short-circuited. In the early stages of acquiring proficiency in a skilled activity, it seems likely that the ventral stream must play a significant role in the visual control of movements.... Conscious perceptual monitoring is coupled with close attentional control over both the overall action itself and the specific limb movements made with respect to the goal object. But as the movements become more and more practised, we would suggest, their visual control becomes encapsulated within the visuomotor modules of the dorsal stream, which together with the pons and cerebellum allow the movements to be executed in an automated fashion. At present, little is known about how this putative shift from ventral to dorsal processing occurs' (2006; 248-249).

There is confusion here. Use of the word 'cognitive' to begin with concerning control, rather than 'conscious', could imply that Milner and Goodale are not proposing that consciousness is involved. However, the word 'conscious' is then introduced. Somehow conscious control becomes 'encapsulated'. But what could this mean? Consciousness is what it is ontologically (supposedly); how could it alter to be purely neural *as* encapsulated, and indeed facilitate automatic movements? What is it that consciousness initially provided that could then be done without? (Neural 'learning' *per se* is quite another story. For what would then happen is purely material transformation, or adaptation.)



This is both implausible as an analogy, and innocent of the difficulties. Firstly (the Malebranche problem), the ventral stream (as the human operator) is the homunculus identifying the object by seeing it and instructing the dorsal stream (the robot). But how does the ventral stream (the human operator) *itself* see the object? *Another* homunculus is required to effect the ventral stream's seeing (another camera and human operator). And so on. We have been through this before, and nothing is explained thereby. Secondly, how can a *seeing* (with its peculiar ontological status) be causally efficacious upon a purely neural activity (the Arnauld problem)?

Milner and Goodale have impacted positively our deliberations. Inadvertently in the C model, they have raised the question of what the brain phenomenon is for, and offered a solution.<sup>58</sup> But as will be shown, that solution is not *via* mentalism.

### 3.3.4 Language

Milner and Goodale have made us sensitive to the distinction between physical processing by the brain, and the appearance of the brain phenomenon. But this distinction is often ignored in neuroscientific literature. Firstly talk is of the brain processing *light* or *colour*. Or we find discussion of how the brain handles *language*. Secondly, it is presupposed that what is *identified* in fMRI scans of the brain is consciousness, be it visual or verbal (or feeling, etc.).

But language as heard, spoken or read is *not* how the brain encounters its input. The brain encounters transduced structured electromagnetic radiation that appears, at some point, to turn into seeing as reading, and transduced structured compression waves that turn into hearing. If there were such a thing as seeing or hearing language, they would be brain phenomena the brain creates upstream (as it were) for a function and as an ontology which, by now in our investigation, have become unspecified, and appear not to lie on a causal path for the organism or person, despite the fact that we talk as if they did. Thus we cannot just suppose the *brain* reads, hears or thinks. Nor in processing input need its areas be visual or auditory. But this is how the literature describes them. We illustrate with two examples.

---

<sup>58</sup> It should be added that the authors do point out that much more information is now available about the complexity of pathways in the brain, but still they endorse their essential model.

### 3.3.4.1 Language and neuroscience

The first is an experimental study reported in 1998, the work of Neville *et al.*, and published by the National Academy of Sciences USA. It is commented upon by Bear *et al.* in 2006 (636-637). Our sole concern is how the data are interpreted by the researchers. We will first outline the study in the researchers' terms.

The *purpose* of the study was to compare the neural activation of three kinds of subject. They were: normal hearing subjects who had no familiarity with American Sign Language (ASL); deaf subjects whose first language was ASL; and normal hearing subjects who were also familiar with ASL because their parents were deaf – i.e. the subjects were bilingual.

The *method* was to present to all groups English sentences on a screen, and ASL sentences provided by a signer on a screen. At the same time, fMRI scans took place on the brains of subjects to determine what locations in the brain were activated. This activation was then to be interpreted for significant differences.

It has been proposed for a hundred years that language areas occur in the left brain hemisphere, the Broca and Wernicke areas (the inferior frontal area and superior temporal sulcus (STS) posterior respectively). Neville *et al.* say:

'Every subject processing their native language (i.e. hearing subjects processing English, deaf subjects processing ASL, hearing native signers processing English, and hearing native signers processing ASL) displayed significant activation within left hemisphere structures classically linked to language processing' (online text: 5).

However, what was also discovered was that for deaf subjects where ASL was the first language, *reading* English (the second language) did not activate the left hemisphere areas. What was activated was the right hemisphere including, and specifically, the superior temporal gyrus – an area activated when hearing subjects *hear* English. In the case of bilingual subjects, reading English activated the left hemisphere areas, and watching ASL also activated right hemisphere areas.

The question is, how do the researchers interpret these findings, and what problems lie within those interpretations? For one confusion is how we might interpret the terms 'seeing', 'hearing' and 'understanding'. All are mentalist terms, but understanding is

differentiable from seeing and hearing. It may be that seeing and hearing presuppose understanding implicitly, but if we associate them more with the senses from which the mentalist terms take their cue (i.e. eyes and ears), rather than the point at which the subject can act responsively (i.e. when understanding, so-called, has taken place), then some clarification of the experimental data might obtain.

1) Although subjects' brains respond to input of their first language in the left hemisphere Broca and Wernicke areas, the subjects' experience (so-called) would appear to be different. A hearing English speaker reading English sentences *hears* them (supposedly) as they read them. But how could an ASL reader hear what they see for, being congenitally deaf, they can have no experience of hearing? Since activation occurs in those areas for both hearing and deaf subjects, the implication is that the processing is *not* as seeing or hearing language, but the ability for *responsive action* apart from the experience (of seeing or hearing).

The researchers do say: 'These results imply that there are strong biological constraints that render these particular brain areas well designed to process linguistic information independently of the structure and modality of the language' (online text: 5). But this already confuses physicality and mentality. For the brain processes structured *electromagnetic radiation* information, not *linguistic information*, or *language* as (supposedly) seen or heard – and *as such* so identified. These must be considered added modalities, currently supposed as mental and manufactured by the brain (somehow). But *does* the brain manufacture those modalities, or in using those terms are we supposing we know something we do not?

2) Then again, what do we infer from the right hemisphere activation? This takes place for ASL readers of English when it is not a first language *and* when viewing ASL. It also occurs for hearing subjects hearing English (according to Bear *et al.*), and for bilingual subjects watching ASL.

Both Neville *et al.* and Bear *et al.* refer to these as *auditory* areas because they are associated with hearing in hearing subjects. Indeed, they both propose these auditory areas are 'recruited' by ASL subjects in viewing ASL. But this is confused terminology resulting from the imposition of a mentalist notion (hearing) on brain areas. The right hemisphere activation, we might propose, involves the processing of the *physicality* of input, whether that is electromagnetic radiation for ASL, or compression waves for spoken English. I.e. the

notion that the areas *are* auditory presupposes that experiential hearing takes place there, which is not established.

3) What we then have to ask is why only the left hemisphere is activated in hearing readers of English, and why reading English by deaf subjects does *not* activate the left hemisphere.

There is a way to address this. If we drop the notions of seeing and hearing, i.e. we do away with the mentalist notions altogether, we can then say that the eye-input of working-ear subjects in the left hemisphere completes the ability for *responsive action* to the input, even if no action takes place actually. (This could be correlated with the notion of understanding, but substitutes it in practice.) The fact that hearing what is seen appears to happen in working-ear subjects does not mean this is to be associated with what comes in through the ears, i.e. compression wave input. For compression wave input that activates the right hemisphere, we might conclude from (2), is pure physical processing and has nothing to do with hearing. What hearing is actually, and when/where it occurs, is a completely separate issue. (In the experiments, the brain-impact of ‘visual’ input has been subtracted from the data. But this is to assume it is not relevant.)

On the other hand, non-working-ear subjects do not *not* hear, for what hearing *is* is not established. Nonetheless, whilst, for these subjects, the eye-input of ASL activates both left and right hemispheres (rendering their ability to act responsively), when the eye-input is for their second language, English, what constitutes their ability to act responsively does not involve the (left hemisphere) areas activated by their first language. According to Neville *et al.*, this could be associated with their learning (written) English later as a second language. This conclusion may be supported by the fact that ‘hearing’ native ASL subjects viewing ASL do show activation in both the left and right hemispheres (though according to Neville *et al.*, lesion studies indicate that the left brain of this kind of subject remains predominant).

4) As stated initially, in these kind of studies it is customary for mental states to be presupposed. Therefore mentality must be accounted for. (Not only do Neville *et al.* not explain seeing or hearing, they do not recognise their potentially obfuscatory use of the terms.) Thus whilst it is normal in one breath to differentiate activation states in the eyes or ears from seeing and hearing as mental states (because they *are* different), in the next breath the states are conflated and layered upon the brain as if they were comprehensible in *only* mentalist terms – and viewable in fMRI scans. (Recall Frith’s words in Chapter 1: ‘By

measuring brain activity, we have an objective confirmation of...mental events.’) This *assumes* we correctly identify seeing or hearing as we (supposedly) experience them – i.e. that we can accurately or veridically introspect our seeing and hearing (as Arnauld, Locke and Block propose).

The mental states of seeing and hearing thus become types of *existents*, as if they were physical objects in the world associated with the objects we e.g. *see* in the seeing (hence the notion of intentionality). These (hypostatized) ‘objects’ are then projected onto the brain and supposed as neurological states in scans. (Recall Locke’s dubious *resemblance* of perceived objects to the objects themselves, thus covertly maintaining in seeing the causality of the objects seen.) But the Slezak discussion has led us into profound suspicion about this kind of approach. For we have not in fact identified what seeing and hearing *are* at all.

An allied and familiar problem is also evident (raised by John Heil). Seeing and hearing are *different* mental states. But if they are both causal, their causality must exist even though the states are different. But how can the brain, which is made of the same kinds of stuff in all areas, *be* different and yet causal as *both* seeing and hearing?

Of course, if our way of explaining what the Neville *et al.* study shows is to say we do not see or hear, but refer instead to the ‘ability to act responsively’, we do not mean our supposed experience is identical to our ‘ability to act’. For it is the word ‘responsively’ that is of prime importance. But that points to something about neural *function* (*not* as experience) which we are not yet in a position to engage.

Finally, Neville *et al.* do not actually tell us what language is. The supposition is that language is some kind of condition which involves users seeing and hearing it, and indeed understanding it. But what does language *do*, and how does it work? These experiments are carried out on a topic whose nature and function is wholly undetermined in the terms of a physical universe. Thus while the results of the experiment are surely interesting, we do not have a scientific expression as to what they relate.

#### **3.3.4.2 Language and action**

The second example is from a text by Jesse Prinz and Andy Clark published in 2004, about the supposed power of thought. In the Abstract, these sentences occur.

‘In the real world, thinking is always and everywhere about doing. The point of having a brain is to guide embodied beings in a complex material world’ (2004; 57).

Significance lies in their example. It is of the well-publicised paralysed stroke patient who learned to move a cursor about on a screen by thinking and/or willing – allegedly (Graham-Rowe: 1998).

An astonishing feat! We knew our thoughts were important, and science fiction has revelled in the possibility that some beings could cause the world to react to their thoughts alone (telekinesis). This well-publicised experiment seems to be a step on that route, with the Hegelian projection that the real is rational and the rational real: a claim originating with the origins of philosophy and pertaining to our divine capacity to think. However, thinking *is* mental and world objects *are* physical.

Prinz and Clark begin their account like this.

‘A paralyzed stroke victim was enabled to control a computer screen cursor using signals from two neural implants. The signals go to an amplifier which relays them to the computer. The patient is then, with effort and practice, able to learn to use their own neural activity to control simple cursor movements. Notice that in order to get the cursor to move, the patient first needs to *experiment with her own motor signals*. The implants, recall, are lodged in the motor cortex – the area controlling bodily movement. The nerves that feed into the electrodes thus carry signals which, normally, would participate in the control of movements such as the raising of an arm, or a leg, or the wiggling of a finger’ (2004; 67).

Apart from the writers’ inappropriate conflation of ‘the patient’ and ‘her own motor signals’, this is acceptable so far. For what appears to be portrayed is operation in a purely material world. But we are suspicious because the electrodes being inserted into the motor cortex means that all the requisite neural processing to cause whatever is going to happen *via* the motor cortex has already happened. And what *is* that processing? Prinz and Clark straightway continue:

‘To successfully move the cursor by thought, the paralyzed patient first tries to will the motion of various bodily parts. When such efforts yield a signal that the computer hears,’ [note that computers *hear* in this scenario] ‘a buzzer sounds so the patient knows to concentrate on “that kind of thought”. After a while, having “that kind of

thought” becomes “second nature” –’ [*habit*, as we have encountered] ‘when you want the cursor to move, you just will it to do so, exactly as you might will your own leg to move’ (2004; 67).

This might fit the popular press, but to claim it is science or philosophy, when neither will or thought are identified in any scientific sense, is greatly inflated. The scientists who implemented the programme doubtless had no interest in whether such things as will or thought exist as properties of the brain. They *should* have an interest, for as we have said from the beginning, without such, neuroscience that depends upon such properties cannot claim to be a science. Prinz and Clark, however, should have an interest, since they are philosophers. It is surely the job of philosophers to question the claims of such an experiment. I.e. has the experiment proved its thesis?

But there is more to this concerning Clark. For he has addressed the work of Milner and Goodale in a paper in 2001. He begins his comments with a quote from Brian O’Shaughnessy (1992).

‘One keeps looking as one guides the finger [to the printed cross], and does so right up until the moment the finger contacts the cross, and the reason, surely, is that sight is continually informing one as to where in one’s visual field to move one’s visible physical finger’ (2001a; 496).

By sight, as Clark says, O’Shaughnessy means conscious visual experience, and it is precisely the assumption that consciousness is involved in doing (or action) that Clark questions, invoking Milner and Goodale. He states that:

‘If the Milner and Goodale story is accepted, there is at least a *prima facie* problem with certain readings of...the assumption of experience-based control. For it is no longer clear that conscious visual experience is deeply linked to the fine-grained control of motor action. The kind of coding and processing implicated in the real-time guidance of delicate, fluent, object-engaged action is, it now seems, frequently and significantly distinct from that which supports our ongoing perceptual experience of the scene. Indeed, the two can often conflict, with little or no effect on the fluent, world-engaged action’ (504).

The obvious question arising from this is: If thought has the same status as vision, both being mental states, why should thought (in the paralysed patient experiment) influence action when vision does not?

Elsewhere Clark appears to contradict his story on Milner and Goodale. He says that:

‘What we see is a world tailor-made for thought, reason and planning. That’s why we can intentionally act in the very world we experience’. Indeed ‘The scene before us is...rich in colour, depth and detail, just as we take it to be. And we have access to this depth and detail as easily as we have access to facts stored in our biological memory’ (2002; 210).

By ‘access’ Clark presumably means the ability to cause action from ‘the scene before us’.

The way Clark seems to resolve the apparent contradiction is to redefine the *role* of (perceptual) consciousness in influencing action. He says:

‘Although it may seem as if our conscious seeing is what continuously and delicately guides our fine-tuned motor activity, such online control may be largely and typically devolved to distinct, nonconscious visual-input-using systems. Conscious visual experience, by contrast, is delicately geared to knowing the visual scene in ways appropriate for planning, recall, and reason based action selection [i.e. thought!]. And it is only in this rather indirect sense that conscious seeing may be said to guide our actions’ (2001a, 511).

(Clark brings to his story the views of Christopher Peacocke and Martin Davies, amongst others.)

But at the heart of this account lies a fundamental issue. Whether or not one proposes to associate perception with reason and thought, rather than ‘fine-tuned’ action, we still have no explanation how thought or perception *could* influence action in any way at all. The Milner and Goodale story (and others of the kind) draw attention to empirical problems which are not resolved by a mere shuffling of function terms in the categories of mentality – whether or not that is a device for *saving* mentality in the face of a lack of empirical proof. For what we are not told is what perception or thought *are* in the physical terms necessary for influencing action. The lack of that prerequisite information precedes, explanatorily, Milner and Goodale’s statement that: ‘We have no real idea what the critical difference is between

neural activity that reaches awareness [i.e. consciousness] and that which does not'. For we have seen, in the experiments of Neville et al., that the existence in the brain of vision or hearing, whether or not they entail understanding, depends *entirely* upon a justification from our own consciousness-derived self-knowledge. And this is, by now, thoroughly dubious.

### 3.3.5 Libet

The work of Benjamin Libet is famous. Frith says:

'Most of the work that scientists do is of little interest outside a very narrow circle of other scientists in the same field.... But occasionally an observation is made that is so startling that it is discussed widely outside the field of science. One such observation was published in 1983 by Benjamin Libet and his colleagues' (2007; 66).

The title of Libet's paper well characterises his findings. 'Time of conscious intention to act in relation to onset of neural activity (readiness potential). The unconscious initiation of a freely voluntary act' (Libet *et al.*: 1983). Libet's findings have been widely discussed and interpreted in many ways. We are not going to explore the literature (which is easily accessible). We make one remark.

In Libet's experiment, subjects were asked to perform a simple task voluntarily. They were asked to 'flex the right hand, fingers or wrist', and report the moment when they were aware of the decision to do so. At the same time their brain activity was recorded by EEG. What was determined was that the recording of the awareness of deciding to flex their fingers/wrist occurred after the brain had apparently already become active in relation to that decision (readiness potential). The time interval was between 350-400 milliseconds (1999).

In terms of consciousness, which is the awareness of the decision to flex the fingers/wrist, it appears the brain has 'already made the decision' before the subject realises it. The obvious conclusion is that we have no free will, because the domain of free will is consciousness, not the material brain.

As said, we will not explore the debates. Our remark is this. If consciousness does not exist, and indeed if something like the C model (§3.3.3) is the case, then Libet's experiments do not problematize consciousness or awareness; they indicate something quite different (assuming they are legitimate). The brain phenomenon misunderstood as consciousness is not

involved in direct causality at all. All that Libet's, and Libet-type experiments show is how the brain operates according to the *function* of the brain phenomenon in fact, which is as yet not identified. In other words, Libet's experiments have been misunderstood, including by Libet himself. That is what our text is moving towards.

### 3.4 Summary

From this chapter, two reasonable conclusions can be drawn.

- 1) The theory of consciousness is incoherent.
- 2) The existence of consciousness is unmotivated.

As a scientific theory, for it is a theory, consciousness is ineffectual. It is neither classifiable nor expressible in scientific terms. It evidently derives from a wrong assumption about how the universe operates, as did the geocentric model in cosmology. In this chapter, which could be written multiple times with different examples, we see the struggle to make sense of consciousness, and the failure to do so. We see that mentalist terms slide from under the researchers' command. Potentially useful findings are ultimately unsynthesizable within one overarching model. That is the incoherence.

But the theory is also unmotivated. The researchers, or theoreticians, rely upon the construct without considering its original plausibility. What is consciousness supposed to do? *Why* are there such things as perception, thought or feeling – in non-tautological explanatory terms?

What is compelling, apparently, is that our internal states *tell us* consciousness exists. But this very attribution is the likely source of the problem. It may be a 'psychological' compulsion, but not a biological one. With biology we know what we are talking about, in the sense that we do not rise above the material world. But upwards from biology, in psychology, we have lost command of our discourse. Our motivation is not scientific, but personal. I.e. from *within* a biology of ourselves over which we have no grasp, for we have no model.

However, the problem with consciousness has not been fully expressed. In the next chapter we get to the heart of the matter.

## The Theory of Brain-Sign

### 4.0 The predicament, & the mentalist error of knowledge

In the previous chapter we explored problems with consciousness, questioning it as a credible account of the brain phenomenon. We now turn to the deeper issue which is consciousness as the provender of knowledge as a cultural and scientific presupposition. First we consider the assumptions in the foundations of knowledge as mental states. Then we review neuroscience and its relation to knowledge. Finally, as a specific example, we will consider the neural network position of Paul Churchland.

#### 4.1 Prelude to the predicament

We talk as if there is a world to see. We speak of history forged by human beings and their conduct, their motives and ambitions. Yet science has discovered another domain, that of a physicality unimbued with such human features, operating under quantifiable laws. (This was the division addressed by Wilhelm Dilthey in the late-nineteenth/early-twentieth centuries, *viz.* *Geisteswissenschaften* by contrast to *Naturwissenschaften* (e.g. 1954).) How can there be a transition from one to the other? For the world as seen (enjoyed, painted, of deep attachment) does not exist if there is no one to see it, or seeing is not possible.

Consciousness has been the supposed transition, for human *being* is founded on the powers of consciousness *via* which scientific description of the universe is generated. But it has proved deeply mysterious that consciousness can effect that transition and yet itself be part of the universe described by science. Even a writer on science, Robert Cease, can say:

‘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.’

For many years the topic has been a concern of Jaegwon Kim. In a recent publication he states:

‘The problem of mental causation is to answer this question: How can the mind exert its causal powers in a world that is fundamentally physical? The problem of consciousness is to answer the following question: How can there be such a thing as consciousness in a physical world, a world consisting of nothing but bits of matter distributed over spacetime behaving in accordance with physical law? As it turns out, the two problems are interconnected’ (2005; 7).

Given our prior exploration, we may find this genuinely surprising. Kim does not begin by considering if the grounds of the questions are correct.

Are there minds? Is a Locke-type delivery of the mind’s constituents credible? Is it reasonable to assume we have direct access to what a mind is? Can there be consciousness when colours present us, not only with the intractable question as to the location of their existence, but whether we actually have understood them (or vision itself) scientifically? How would consciousness be causal when empirical research indicates that it does not lie on a causal path, neurally? Why does the issue of what perception, thought or feeling *are* not precede working out whether they could be reduced to the physical universe? And so on.

Here is Kim on non-reductive physicalism. ‘There is no consensus on exactly how nonreductive physicalism is to be formulated, for the simple reason that there is no consensus about either how physicalism is to be formulated or how we should understand reduction’ (2005; 33).

Kim does not ask whether *mentalism* can be formulated. The one thing we may be sure of, for we live in a world in which it is obviously causally operative, is physicality. But it is physicality Kim questions. Indeed, he actually says: ‘For most of us, there is no need to belabour the centrality of consciousness to our conception of ourselves as creatures with minds.... Consciousness research is thriving’ (10-11).<sup>59</sup>

Research may be thriving, but if its target is not viable, it has negligible likelihood of results. Kim’s expressed position, that ‘the centrality of consciousness to our conception of ourselves as creatures with minds’, while likely correct, fails to address the issue. For it does not identify what consciousness *is* ‘for most of us’. Little wonder it is so difficult to locate in

---

<sup>59</sup> To be accurate to Kim, the whole of this last sentence reads: ‘Although consciousness research is thriving, much of cognitive science seems still in the grip of what may be called methodological epiphenomenalism’ – which undermines his confidence that consciousness research is significant.

the physical universe. From the previous chapter, it appears incapable of description in a way acceptable by all.

Where does the question of consciousness begin? Kim says it begins with our ‘conception of ourselves’. This is surely mistaken if we want to characterise consciousness scientifically. The questions should be: What causes the brain phenomenon assumed as consciousness, why, and what job does it do biologically? Kim’s approach already assumes what needs to be investigated.

The question is even more personal. What happens in the physical universe when a philosopher makes a statement? What is a statement in bio-physical terms?

What Kim presumably means by ‘most of us’ is ‘what *I* want to talk about as *I* think about it’. But what *is* Kim’s thinking? For a scientific explanation, Kim must submit himself and his (so-called) thoughts to evaluation as a biological organism. For like most of the philosophical community, he takes thinking for granted. How could ‘bits of matter’ think? as Kim disparagingly refers to physicality. On the contrary, matter can be highly organised, whether it thinks or not.

#### **4.1.1 The predicament exposed**

John Heil says: ‘Presumably we have something like direct access to ways we think and talk about the world.’ This occurs on page 1 of his book *From an Ontological Point of View* (2003). The notion, which pervades philosophy from before, through and beyond Descartes, is likely the fundamental mistake. (Bishop Berkeley is its apotheosis: not only can we know only the contents of our own minds, what exists *depends* upon its being known, specifically seen (if only by God). Which led to Dr Johnson’s physicalist foot as he kicked the stone: ‘I refute it *thus*’.)

Why do we say direct access is the fundamental mistake? Because it assumes what it has not demonstrated. ‘I know the world because I can know myself knowing the world’ is not the place to begin. What is it to *know* the world? Philosophy’s most pervasive answer is representation of what is to a subject.<sup>60</sup> But as we have seen, this is not a convincing answer,

---

<sup>60</sup> There are alternative scenarios: particularly influential is that by Heidegger in *Being and Time*, e.g. Being-in-the-world and Being-with with others. We shall come to this later.

for its terms are not defined or explained beyond assertion. Assuming the universe is purely physical: What is a (neural) representation? What is a (neural) subject? What is (neural) representation to a subject? Why is this knowledge? How could mental structure, particularly the subject/object relationship, have causal properties?

Yet without establishing a physical/causal premise for knowledge (i.e. without solving the *Geisteswissenschaften/Naturwissenschaften* dichotomy), philosophy further claims we know we know, *thereby* justifying knowledge *per se*. ‘Presumably we have something like direct access to ways we think and talk about the world.’ This states there is such a thing as thinking and (by some unspecified means) we have ‘something like’ direct access to it.

What is direct access? That we hear/see our selves think?<sup>61</sup> But again, what is thinking or seeing or hearing? José Bermúdez says: ‘Our introspective awareness of ourselves is fundamentally different from our awareness of ordinary physical objects and other psychological subjects’ (2003; 213). But while a *difference* might be conceded, he does not explain what the latter are or how they exist from which the former should be distinguished. One cannot just decide by *fiat* that biology gives us an actuality of awareness, etc., and direct access to it. This avoids the central issue, though it characterises the history as we have seen.

#### 4.1.2 The predicament considered

Suppose the brain created such a thing as thought. Let us say it is represented by the following sentences, about a walk at dusk along the promenade by the seaside.

‘I remember, from when I was a boy, seeing these gaily coloured lights strung from the lampposts. I remember it gave me an excited sense, a joyous sense, an – alas, I can’t quite explain what the mysterious sense was.’

This thought is associated with a seeing and a feeling. Of course I am (supposedly) aware of this thought. But do I *know* what it is?

A standard response could be: Well, it’s the thought *that* ‘I remember, from when I was a boy’, etc. But identifying thoughts as propositional attitudes will not help. Dennett and

---

<sup>61</sup> Differentiation is made between seeing the world and ‘seeing’ access to mental states, e.g. with Sydney Shoemaker (1988). But our investigation is more primary than that.

Churchland have addressed propositional attitudes critically:<sup>62</sup> certainly they offer no ontological insight. What we are asking is: Do we know what this thought is *qua* existent? Or the seeing or feeling?<sup>63</sup>

What I see are coloured lights; but I have no knowledge of what my seeing the coloured lights *is*, nor what the thought about them *is*. I cannot get *before* or *behind* the seeing or thinking or feeling to make of *them* a representation to a subject, if that is what is supposed to generate knowledge. So, *pace* Heil, I have no capacity to know what a seeing or thought is because I have no mechanism to render them apparent as a knowledge of them (by bringing them to consciousness). The seeing (so-called) is what I *am*, not what I can know. The thought is what I *am*, not what I can know.<sup>64</sup>

Fred Dretske has attempted to tackle this problem. He says: ‘I can be made aware (of the fact) that I am stepping on an ant by actually seeing myself step on an ant (an event) but that is not the way I become aware that I see the ant. I don’t see myself see an ant’ (2003; 8). Dretske is searching for a means to verify consciousness exists. Thus he says: ‘I’m not asking *whether* you know you are not a zombie. Of course you do. I’m asking *how* you know it’ (1). But here and in subsequent texts he fails to distinguish two different questions, because (as he says) he does not doubt we are conscious.

1) By what means can I tell I am conscious?

2) Why are those the means to know I *am* conscious?

Thus he proposes to answer question 1 but, because he does not pose the problem comprehensively, he fails to address question 2 or even recognise its significance. For clearly if one asks the question: How do I know I am conscious? one already assumes one is conscious – which invalidates the inquiry, for then one merely finds the means for what one assumes to do the job. The question Why? is not truly posed.

If the brain has made what we term a seeing or thought for its purposes (whatever they may be), there seems little reason to suppose it can render what the seeing or thought are to my knowledge, because my so-called knowledge (as seeing or thought) serves its (the

---

<sup>62</sup> Dennett says: ‘The most sweeping conclusion I have drawn...is that the large and well-regarded literature on propositional attitudes...is largely a disciplinary artefact of no importance whatsoever, except perhaps as history’s most slowly unwinding unintended *reductio ad absurdum*’ (1994; 241). Also Churchland (1981).

<sup>63</sup> Obviously we cannot appeal to a Kantian *inner sense*, for it offers no explanation of itself.

<sup>64</sup> This formulation is expressed by Sartre in his text *Being and Nothingness* (1957), although his conclusions and ours are different. Hence his subtitle: *An Essay in Phenomenological Ontology*.

brain's) purposes, not mine (whatever that may be) *as* knowledge. (Recall the discussion at the end of §3.1.5.) That is, I may *seem* to see the coloured lights, but I cannot confirm that the presence to me of the lights results from my *seeing* them.

Put more generally: How could a physical state ('the "dark" electro-chemical brain operating without a grasp of anything') present to itself *knowledge* of its constituents *which are themselves this knowledge*? How could a being (consciousness) obtain the knowledge of its being-creator out of itself as created? This appears more improbable than pulling oneself off the ground by one's own boot laces. Little wonder consciousness has been regarded not merely as mysterious but miraculous. So it is that we are led to the conclusion that consciousness must be considered a mere *theory* of the brain phenomenon, not an indubitable fact, for we have no means to verify its existence *via* itself as the only means of knowing anything (supposedly). This is the critical response to (Dretske's) question 2.

Brentano's characterisation of mental states as *intentional* does not illuminate the situation. Calling them intentional and rendering them problematic *vis-à-vis* the physical offers no route to an explanation concerning the *brain's* activity, because such characterisation adds nothing to our questioning.

If we now return to the inheritance of (knowing) mentality from God, we see that Descartes, Malebranche and Arnauld understood mentality, albeit in somewhat different ways, such that knowing offers transparent access to the universe in the same way that God has it (in principle, if not in fact) *because* he made the universe and knew everything there was to know about it, including himself. This was their prescientific interpretation of their experiential condition.

The idealists' programme presupposed: If it is impossible for us to know *everything* (as God), wherefrom comes our knowledge of *anything*? We know because, in principle, we know what we know *via* the (he who knows all) God analogy. (For Malebranche we know 'in God'. For Berkeley it is God's vision that holds the immaterial world in being if otherwise unseen.) In this vein Block's rational control of action (§3.3.1) means that access-consciousness gives us what *is*, and as such, *consciously*, we can act rationally. But (bypassing Locke *et al.*'s lack of explanation), since we are made from clay, or 'bits of matter', it is wrong to equate what the originators of this discourse meant by knowledge (in its inner realm) with what a physical object can be or do – a distinction Block fails to make. In fact we have *no* explanation, in physical terms, of where knowledge comes from (since it

does not come from God). We can *say* we know, and we can *sense* we know, but *knowing* in philosophical history (which for us enquirers has become *merely* another fact to know) means something which is utterly obscure. It has been bequeathed to us through our culture as an invisible impossibility.

We have unveiled a mystery. It is not consciousness, and how it emerges from ‘bits of matter’. This formulation is really a terminus in current philosophy. The mystery is: Apparently I am I-know-not-what, and I have no idea how to proceed on identifying what I am. We refer to this as our *predicament*: the predicament of physical organisms, given that it has been supposed (without scientific justification) they have what is referred to as knowledge. The predicament is the stumbling block upon which knowledge founders.

The persistent questioner might say, ‘Of course I know that I know. I have a pain in my foot. I know I have that pain. You can’t tell me otherwise.’

But the response is the same. To know one has a pain in one’s foot requires we know what knowing what a pain in one’s foot is. But we cannot get before or behind the knowing of the pain in one’s foot to know it *qua* existent (cf. Dretske’s failure to address question 2). Surely we may acknowledge a differentiation between the pain in one’s foot and (appearing to) know it – indeed this may have significance once we have a viable bio-functional model. But the predicament still exists for knowledge, whether pre-reflective or reflective (so-called).

There is no denial here that we appear to know, i.e. we appear to be aware of what it is that causes us to act. This ‘what we appear to be aware of’ is the *sense* that we know. What we need to get a grasp of is what causes us to have such a sense, and what biological role it performs. We cannot just assume it does what we think it does, for our ‘thinking it does’ is also a brain product, the significance (and ontology) of which is yet to be established. We must start at the beginning, not half way through.

The predicament, therefore, is the condition in which we ‘see’ that not only do we not know, we have (as yet) no theory of what the brain phenomenon ‘we are’ is in the physical universe, which is not knowledge.

### 4.1.3 The predicament and free will

There is, however, a complementary aspect to this. Since we have no idea what will occur to us next *as* (so-called) consciousness, why are we not continually astonished at what does occur? This astonishment should be totally incapacitating. Every (supposed) sight, thought, feeling is a complete novelty in its occurrence (even if we have seen, thought, felt it before). It is just *there*, and ‘we’ are there with it. (We may call this, with deference, the existential condition). We cannot alter it: it has already happened. Could this really be the *freedom* consciousness endows?

The answer appears to be that our being as the brain phenomenon (normally) is *unattuned* to the fact that we have no idea what will occur *as us* next. The implication is that the brain phenomenon is a controlled state which does not serve the purpose of knowledge which *would* continually take us by surprise, but rather that the encounter with the world in its continual originality is elsewhere in our organism, and the brain phenomenon we are is a derivative state serving another function than knowledge. And it is an utterly determined function.

Of course, the notion of a completely determined (functional) state which we are implies we are not free. For if the state is determined, then *we* are determined. This is not novel. Schopenhauer put it in his prize winning essay of 1838.

‘Man does at all times only what he wills, and yet he does this necessarily. But this is due to the fact that he already *is* what he wills. For from that which he *is*, there follows of necessity everything that he, at any time, *does*. If we consider his behaviour objectively, i.e. from the outside, we shall be bound to recognize that, like the behaviour of every natural being, it must be subject to the laws of causality in all its severity. Subjectively, however, everyone feels that he does only what he wills. But this merely means that his activity is a pure expression of his very own being. Every natural being, even the lowest, would feel the same, if it *could* feel’ (1960; 98-99).

In 2002 Daniel Wegner says:

‘We feel that we consciously cause what we do; and yes, our actions happen to us.... Rather than a ghost in the machine, the experience of conscious will is a feeling that helps us to appreciate and remember our authorship of the things our minds and bodies do’ (ix).

Wegner claims he is offering a new account in the free will debate. But Schopenhauer said, in explaining his position 163 years earlier: ‘The consciousness of self-determination and originality which undeniably accompanies all our acts...is therefore not deceptive.... From what we do we know what we are’ (1960; 98).

But comparisons apart, the problem with these explanations is that they obscure the issue. In denying we are free *via* our conscious will (which would appear to be the case), why *should* we nevertheless *know* what we do? If our acts are determined by our physical being, why does our physical being then create another state (consciousness) so that *we* (who?) know of its causality? These ‘explanations’ are simply a means to try to accommodate irreconcilable (apparent) facts.<sup>65</sup>

Moreover, they still leave us with inexplicable knowledge, even if its function is not to be causal upon action. How is this knowledge of ourselves and our actions supposed to *be* in the physical universe?

## 4.2 Bennett & Hacker, and the ontological issue in neuroscience

Having identified that knowledge is not well-defined, and indeed in terms of the early idealists, is wholly opaque (though specified), we must now consider how neuroscience, as a science of matter, addresses knowledge.

The proposal that consciousness (the *only* domain of knowledge) emerges from matter, without consideration of whether such a state actually exists is, as we have said, an indication of a terminus in philosophy *and indeed* neuroscience.

We have already referred to the recent book by Maxwell Bennett and Peter Hacker. Their work arose from the conviction that ‘not all was well with general theorizing’ in neuroscience (2003; 1). Our position certainly approves this thesis, and Dennett and Searle, who have taken part in dialogue on the book’s content, do agree some of the points made (Bennett *et al.* 2007). However, the Bennett and Hacker text is dominated by both mentalism, and the significance of language (the Ordinary Language Philosophy of the 1960s, to which Dennett refers (2007; 73)), neither of which are penetrated in an adequate ontological way.

---

<sup>65</sup> This criticism applies to J.J. Gibson’s *affordances* (1977), and Alva Noë’s *Action in Perception* (2005).

Currently neuroscience, in any global sense – i.e. as dealing with the totality of the relevant phenomena – is impossible. Bennett and Hacker put it thus: ‘[Our] unease was produced by a suspicion that [in neuroscience] in some cases concepts were misconstrued, or misapplied, or stretched beyond their defining conditions of application’ (2003; 1). We have seen this in the text on Neville *et al.* (§3.3.4.1) and their work on deaf and hearing subjects.

Bennett and Hacker state:

‘Leading figures of the first two generations of modern brain-neuroscientists were fundamentally Cartesian. Like Descartes, they distinguished the mind from the brain, and ascribed psychological attributes to the mind. The ascription of such predicates to human beings was, therefore, derivative – as in Cartesian metaphysics. The third generation of neuroscientists, however, repudiated the dualism of their teachers. In the course of explaining the possession of psychological attributes to human beings, they ascribed such attributes not to the mind but to the brain or parts of the brain’ (2003; 68).

It seems that neuroscientists, losing patience with philosophers and their quibbles, decided to ignore the problems and proceed with a vocabulary that lost touch with meaningful reference.

Bennett and Hacker give a good account of the situation that has ensued. We will choose but one example that they do not. This occurs in the book by William Bechtel and Adele Abrahamsen on connectionism (purely physical states) and the mind (2002). They propose that ‘connectionist networks are adept at recognizing patterns’ (282) and they talk of ‘confronting the practical problems in making a robot perceive’ (283).

Connectionist networks do not recognise and robots do not perceive. Only humans beings recognise and perceive because they are conscious, for consciousness is how recognition and perception occur. (We have already encountered Prinz and Clark referring to a computer ‘hearing’ in their discussion on the paralyzed patient.)

It might seem a trivial matter, nit-picking indeed, to point out these misuses of language. The point is, however, that they obscure the crucial scientific issue which Beaumont refers to as ‘neuropsychology...in a conceptual morass’. For they assume a simple parallel between what conscious human beings do, and what machines or technology can do. Bechtel and Abrahamsen have no solution to the problem of consciousness: they certainly do not demonstrate that such a state of the brain exists. So why should they assume that

connectionist networks or robots can do what it has not been demonstrated human beings do? The obscure is made even more obscure. Thus our criticism is not that mentalist language should be preserved for conscious beings. It is that neuroscience is pushed further into the dark by a vocabulary that is already unviable, and certainly does not relate to what physicality, be it machine or brain, actually does.

Dennett responds to Bennett and Hacker. But his distinction between the (neural) *subpersonal* and their hierarchically massed occurrence as *personal* (which as a neural product in his ‘virtual machine’ analogy is consciousness)<sup>66</sup> fails either to justify or explain his position. Indeed, he explicitly disclaims responsibility for the implausibility of the distinction. In one place he says:

‘One may be tempted to ask: are the subpersonal components real intentional systems? At what point in the diminution of prowess as we descend to simple neurons does real intentionality disappear? Don’t ask.’

And conversely:

‘At what point in evolutionary history did reason-appreciators, real selves, make their appearance? Don’t ask’ (1994; 240).

In other words, not only does Dennett assume mentalism without enquiring whether such exists which the brain can effect *personally*, he invites silence on his equation of massed *subpersonalism* to the (mentalist) *personal*. I.e., it is not that no hierarchical operative neural structure could exist: it is that neural consciousness *as* personalisation is assumed without explanation or definition of consciousness itself. Why (and how) is it *this* consequent phenomenon? (Dennett’s strategy, as Seager says, is not to explain consciousness but to ‘explain it *away*’ (1999; 118).

Dennett quotes Wittgenstein as Bennett and Hacker’s source for the mereological fallacy they use to criticise neuroscientific language. Here it is.

---

<sup>66</sup> Dennett attempts to avoid his original notion that consciousness arises when the subpersonal crosses a threshold to the personal, by proposing there is no such ‘finish’ line (1991). But this still does not explain the rise from the subpersonal to the personal which occurs as consciousness in a neural *judgement*, as critics have pointed out.

‘It comes to this: Only of a human being and what resembles (behaves like) a living human being can one say: it has sensations; it sees; is blind; hears, is deaf: is conscious or unconscious’ (2007; 78) (Wittgenstein: 1967; para. 281).

Dennett quotes himself (from 1980) in approving this.

‘I understand English, my brain doesn’t – nor, more particularly, does the proper part of it (if such can be isolated) that operates to “process” incoming sentences and to execute my speech act intentions’ (2007; 77).

Indeed, regardless of his failure to explain how the massed subpersonal turns into the personal, Dennett now proposes to extend Wittgenstein’s expression by noting that robots and chess-playing computers ‘and, yes, brains and their parts do “resemble a living human being (by behaving like a human being)” – and that resemblance is sufficient to warrant an adjusted use of psychological vocabulary to characterize that behaviour’ (78).

Thus the issue is not only laid before us, it has a specious justification. For precisely the opposite conclusion is to be drawn from Wittgenstein’s behaviourist supposition. There is no justification for ascribing psychological language to a robot because it (supposedly) behaves like a human being, for the issue is not behaviour but whether it *has* ‘sensations or is conscious’, i.e. possesses the *psyche* that causes its behaviour. For no investigation of what sensations or consciousness *are* has preceded such claims. Without such, the notion that consciousness exists, with its knowledge property that presumably Dennett depends upon because ‘I understand English, my brain doesn’t’ is merely assumed. Wittgenstein is plainly wrong to propose behaviour as the arbiter of whether an organism (or bit of same, or robot, etc.) is conscious or not. Whether an organism is conscious is a scientific issue, and depends upon ontological theorising and experimental investigation of the assumed state.

Moreover, returning to our prior discussion of Kim’s statements, we have no idea what language, in which claims are made, *is*, or how it operates. (What does Dennett mean by ‘I understand language’? What is ‘understanding’ supposed to be, and who or what am I?) Language must be subject to ontological analysis and experimental investigation as a phenomenon *in the physical world*. But we have seen in the case of Neville *et al.* that language as supposedly *heard* or *seen* is not on the horizon of investigation (though something is widely studied *as* language). And this is because the nature of the problem has not been recognised, by Bennett and Hacker, by Wittgenstein or Dennett, or Neville *et al.*

To return to Bechtel and Abrahamsen and robots. In saying robots perceive, what is lost is the relation of the robot to the world which involves the modification of its physical states for the target object. The scientific issue is the *method* of identification by the robot of the object – *viz.* how does this happen *via* its physical states through sensor input to internal modification? – with the resulting action if relevant. The robot's modified physicality has nought to do with a (robot) subject having a represented mental object of what is in the world *and thereby perceptual knowledge*.

And crucial to the method of identification is the function-orientation of the robot to the object. Is it to avoid it? Is to interact with it in some way? Thus how does the robot (in purely physical terms) establish the precision of this function identification? The physical modifications to achieve this kind of target are what makes robot design challenging.

Indeed the Milner and Goodale example (§3.3.3) of a robot on the moon guided by a human subject *exactly* points to the limits of robot design: how much more effective it would be in solving the problem of object identification and response if a conscious human supported it! Why? Because a conscious human subject has exploitable *knowledge* (supposedly) which the robot has not. This knowledge, hypothetically, is *not* orientated to purpose, but is available for *any* purpose because what is so about the universe (in its objects and states, as God created them) is revealed by conscious perception (Stich's inferential promiscuity, Block's access-consciousness).

But knowledge, so considered, has proved intractable not only to existence, but intelligible specification. In the absence of knowledge, we may now conjecture that the same kind of analysis for a robot is necessary for the human brain. What kind of neural modifications take place in relation to an identified object, and how is there resulting action? To introduce the notion of *perception* for a human brain as neuroscience does, requires an explanation of the perceiving state, one in which a mental subject has an object represented (cf. the topic of colour, §3.2.1.). And that is distinctly improbable.

The Bennett and Hacker approach, by contrast, is to say (with a Wittgensteinian flourish) it is the *person* who perceives, not a brain. But this requires an explanation of a person, which they do not offer (as Searle rightly says). Indeed, the claim that a person perceives depends, not upon a usage of vocabulary, but on a biological *account* of perception, and a person, which Bennett and Hacker do not offer. Thus they simply substitute one fog for another.

### 4.3 Paul Churchland's connectionist knowledge

Both Churchland and Dennett have bemoaned their 'substantial isolation' or 'loneliness' as protagonists of a dynamic physicalist stance for consciousness, against the majority approach in which the significance of consciousness is its state or content – which thus may be non-physical (*qualia*) or irreducible (intentionality).<sup>67</sup> Whilst acknowledging each other, they do not agree. Does Churchland's physicalism offer any insight into the nature of knowledge itself?

Because Churchland proposes causality in the brain is analogous to the connectionist neural network model, and this is widely recognised as closer to (though not identical with) actual neural structures than the serial digital computer, we are more sympathetic to his position than Dennett's. But his account of consciousness (or experience) involves no justification for its existence in the first place. While *we* will identify an actual physical state which the brain phenomenon could be, Churchland accepts consciousness as philosophical history supposes. He claims that to be embraced by science, all that is needed is the correct intertheoretic reductive analogy.

For example he states that: 'the brain is a medium selected for its ability to assume, hold, and deploy the conceptual systems we call theories.... A theory is the brain's way of making sense of the world in which it lives, an activity that is its original and primary function' (2007; 5).

It is immediately evident that a theory is to be equated with a state of the brain. This condition *is* a consciously held theory. It is not a written sentence of the theory (or propositional attitude), but the specific neural condition of the brain.<sup>68</sup>

For Churchland, the consciousness of a theory is claimed to be directly reducible to the condition of the neural network on the model of intertheoretic reduction. Of course the theory will not be permanently conscious, because other states will 'become' conscious. Churchland offers an account of this.

---

<sup>67</sup> See Churchland (2007; 1-2), Dennett (1994; 243).

<sup>68</sup> E.g. 'A given concept encompasses a substantial *range* of distinct but closely related cases, in that the mature network will have generated, in the course of its learning the concept, a proprietary *volume* within its activation space, a volume that confines all of the possible points (i.e. neuronal activation vectors) that count as determinate cases of the abstract determinable that is the concept proper' (2007; 144).

‘Such an attentional focus [as e.g. this theory] is also movable, from one learned category to another, as a function of the network’s unfolding activation patterns or “frame of mind” at its higher neuronal layers. Such a network has an ongoing *control* of its topic selections from, and its conceptual interpretations of, its unfolding perceptual inputs. In particular, such a network can bring to bear, now in a more selective way, the general background knowledge embodied more or less permanently in the configuration of its myriad synaptic connections’ (14).

(Churchland’s use of the expression ‘synaptic connections’ is a supposition that neural networks simulate what synaptic connections do. Note, too, the equation of ‘activation patterns’ of the network with a mental state, or ‘frame of mind’.)

A prime question in neural science is, of course, how the brain actually operates to effect its control of the body in the world. The problem with Churchland’s account is not that it does not contribute to the literature of how the brain might work, but why it has anything to do with consciousness – or a derivative knowledge. Churchland specifically refers to ‘background knowledge’ encoded in the ‘myriad synaptic connections’. His is a *cognitive* account, one in which the function of the brain is to develop and hold theories in its structure which will influence what the organism does. But this kind of account is already hostage to the notion of consciousness as an entirely different ontological order from ‘bits of matter’, one historically derived from God’s knowledgeable creation of all that is in the universe. But, since the account demands to be evolutionary and biological, it must force analogies and make suppositions which are both beyond plausibility, and are non-explanatory.

While Dennett proposes subpersonal states accumulate to become personal (and consequently conscious), Churchland equivalently proposes that networks achieve the ‘sculpted space’ of ‘attractors’ or ‘prototype wells’.

‘Sculpted space is the conceptual framework of that layer of neurons’ (13). ‘A recurrent network can represent the changing world with a continuous *sequence* of activation patterns at its second layer’ (14). Indeed, ‘reliable prediction [of the system] becomes utterly impossible.... The system is too mercurial to permit the prediction of absolutely specific behaviours at any point in the nonimmediate future. Thus emerges the spontaneity we expect of, and prize in, a normal stream of conscious cognitive activity’ (15).

(Note that for Churchland, ‘the spontaneity we expect...in a normal stream of consciousness’ is not astonishing and incapacitating, as we conjectured, but ‘expected’ and ‘prized’. Frankly this is Panglossian, and avoids any penetration of the extraordinary nature of the circumstance.)

But why is all this *network activity consciousness*? What Churchland would have to do to establish his proposal is to identify what knowledge is in the first place, and how would he do that? We have no way of getting at it so to classify it. In the universe, in fact, it would seem we exist without knowledge (not least, because we have no idea what it is). And given the brain is ‘bits of matter’, that is entirely in keeping. This does not say, in principle, that attractors and sequenced activation patterns have no place in a causal account of the organism. But in proposing that particular network states are consciousness, we lack any sense of how the two kinds could correspond (how one could ‘emerge’ from the other, as Churchland says), let alone how a reduction could take place.

Churchland depends upon an analogy to intertheoretical reductions in science. Examples are temperature and mean molecular kinetic energy, and light and electromagnetic radiation. Thus temperature *is* mean molecular kinetic energy. Light *is* electromagnetic radiation. Because such reductions have been achieved, the difficulty with colours can be similarly addressed.<sup>69</sup> Colours *will be* brain states.

But this is an entirely improper approach, because Churchland’s exemplified reductions depend upon circumstances which he does not investigate in a valid ontological way. It may be that temperature, as we experience it, is correlated with mean molecular kinetic energy *in that* the temperature of a body can be measured by an instrument and associated with mean molecular kinetic energy. But why do we *feel* a body to be hot or cold, and how do the *concepts* of hot and cold come into being? More specifically, why do we *know* that we feel a body to be hot or cold? Surely we may suppose that receptors in the skin have their signals transposed to conditions of neural networks in the brain, and these conditions may cause us to remain sitting by the fire, or come in out of the cold, as pure physical actions. But how do we do this because we *know* we are having the requisite experiences?

Churchland’s implicit contention is that our knowing experience *is* a form of physicality. And this contention is as mysterious as Levine’s supposition that he can

---

<sup>69</sup> Churchland acknowledges his inspiration to Mohan Matthen (2007; 198), whose work we have encountered.

problematize the property of *qualia* because he has access to it (in a *bona fide* way) as an objective state *via* his subjectivity. The problem is the same in both cases. For neither Levine nor Churchland identify what knowing experience is, but each claims for it, in opposite directions, what he requires for his theory without attempting to place it *first* in the universe *in terms we already identify as physical*. Churchland's use of the analogy with light and electromagnetic radiation is deceptive. Concerning colour and its equation with reflectance properties ('efficiencies' Churchland terms them):

'On the modern reductive approach, this is no more surprising...than the identification of light itself with electromagnetic waves. Such identities may surprise the scientifically uninformed, but they leave the objective reality of light...entirely intact' (2007; 199).

But *light* (as we have considered) is both a mentalist condition and mentalist notion (i.e. what we notionally see) projected onto 'objective reality'. In the universe, science informs us, there is electromagnetic radiation, not light. It is precisely the apparent fact that we see light (or colour) that Churchland *first* has to explain. He cannot just claim a reduction of the objective (electromagnetic radiation) to mind or consciousness (light) and then *vice versa* consciousness (light) to the physical world (neural network states) on a supposed scientific analogy from the physical world. This is no better than Dennett's 'Don't ask'. For it covertly operates by the double reduction displayed, from the physical to the mental then claiming the mental is the physical, but assuming the scientific reduction is but one reduction, physical to physical. Indeed, Churchland 'addresses' precisely this point.

'To ask for an explanation of why a given qualia is 'correlated' with a given activation vector [of a neural network] is to ask for some natural law or laws that somehow connect qualia with activation vectors of the relevant kind. But there can be such a natural law only if the quale and the vector are distinct things..... In the case at issue, however, the proposal is that the qualia and the vectors are not distinct things at all: they are identical...although known to us by two different names' (195).

But the issue is *not* that they are 'known to us by...different names'. The issue is that ontologically they appear to be of different kinds. Yet both depend conceptually upon *a* (mental) kind, *viz.* consciousness, which itself is not explained by simply defining it as the other kind, physicality. How is it that the second kind, physicality (and its scientific reductions), gets to be known by consciousness *as* the second kind? Certainly not by analogy

with light and electromagnetic radiation, both of which are (supposedly): (1) of the second kind, physicality, but only apparent (2) by the first kind, consciousness. Nothing we need to know scientifically about *consciousness* has been explained. For knowledge itself, which we *are* (apparently), remains unspecified. And as we have shown *via* the predicament, it is not available to 'knowledge' to be known as knowledge – whatever that may be.

So, for example, if we built a model of the brain's neural network structure on a computer, or in other than biological means, would it be capable of experience and knowledge? The answer can only be: 'Don't ask'.

Why does Churchland get himself into this position? It is the stance from which he begins, the assumed stance of the notion of mind. He commences by claiming: 'the brain is a medium selected for its ability to assume, hold, and deploy the conceptual systems we call theories.... A theory is the brain's way of making sense of the world in which it lives, an activity that is its original and primary function.'

This claim exists to support a pre-designated position (similar to Dretske earlier). If we take seriously Darwin's evolutionary proposal, then the function of the organism is to survive and reproduce, nothing more. To survive and reproduce, organisms, by natural selection, adapt, and apparently this can become increasingly elaborate in new species evolving. One organ that certainly has significantly developed under evolutionary principles is the brain. But the brain's function is 'no more than' to maintain the organism in its survival and reproduction. One might well suppose that a method of achieving this is the development of neural network-type capabilities. But the physicalities of a neural network are not theories, in the terms in which we understand theories (i.e. as consciously held states). They are and remain a physical adaptational modality which could be causal for the organism. What else happens with the brain, deriving *from* the causal processes of the neural network (and other brain operational modi), is as yet to be determined.

Churchland's use of the notion that 'a theory is the brain's way of making sense of the world in which it lives' assumes the brain is *capable* of 'making sense'. But making sense is what a person or mind does (notionally, or as Bennett and Hacker would say). All a brain can do, but obviously it does it extremely well as mere 'bits of matter', is alter its physical states and be *thence* causal.

Churchland is not ignorant of his usage here. He is quite explicit about his aim, and it accounts for the position he adopts. *We* have dispensed with the notions of both knowledge and truth. But Churchland says:

‘Our model here might be that fringe account of classical truth occasionally advanced by the pragmatists, wherein it was proposed to define the truth of any representation directly in terms of behavioral or navigational success to which it gives rise: crudely, a true proposition is *one that works*. In the same spirit – but by-passing beliefs, propositions, and truth entirely – we might attempt to define any representation as an instance of genuine *knowledge* just in case, when deployed, it produces successful behaviour or navigation. This tempts me hardly at all. While I am at least a closet pragmatist, and while I resonate to the idea that the brain’s cognitive activities are ultimately in the service of motor control [i.e. action], I balk at any such direct definition of the sort proposed.... Specifically, if we *define* or *identify* what counts as truth, or as knowledge, in terms of the behavioral success it produces, then we will not be able to give a nontrivial *explanation* of those behavioral successes in terms of cognitive representations that do give rise to the level of knowledge and truth’ (103).

This mission statement could not be more clear. Unhappily it undermines Churchland’s avowed aim: to provide a physical account of the organism in the world. For he is constrained by mentalist suppositions, as is Dennett.

‘One’s perceptual resources are only as good as the peculiar background conceptual framework in which, for each individual, those representations are destined to be expressed. And such a framework is always and ever fragile, an imperfect, and, in the long run, an *ephemeral* attempt to grasp the structure of the world. Plato’s timeless heaven is, in the end, a poor metaphor for one’s general conceptual framework’ (112).

Plato’s heaven may be an ancient myth. But why does Churchland not explore the notion, deriving precisely from this statement, that conception *itself* as a cause of behaviour is a myth? After all, he has laboured to accumulate a degree of physical structure in his neural network descriptions which could account (at least in part) for behaviour.

But what is important for the philosopher, despite his emphasis on science and scientific method, is that he can revert to the mind’s control of the world (our experience as such), even if its efforts are ephemeral. This control is confusedly mixed up with the

philosopher's professional need to *explain* the world, for which these same mentalist (control) categories, as *terms* of the explanation, are required. In other words, Churchland's process of explanation about the world is positioned outside the world, which is of course impossible.

This is unfortunate, for Churchland's labours are in many ways informative, as we shall come to later.

#### 4.4 Conclusion

We are now in a position to develop and appreciate the benefits of brain-sign theory. To get there we have had to demonstrate the inadequacy of previous interpretations of the brain phenomenon as consciousness. It is the structure and operation of the biology of the organism that we must get at to see why the misinterpretation has taken place – or indeed why *any* interpretation takes place. However, to do that we must set aside our own conceptual saturation in the supposition that we exist as consciousness, and this is very difficult *because* of our biology. We must appreciate that we do not rise out of biology as consciousness, but remain entirely within it. Even biologists are culturally trapped. Here is Richard Dawkins:

'Each one of us knows, from the evidence of our own introspection, that, at least in one modern survival machine...purposiveness has evolved the property we call "consciousness". I am not philosopher enough to discuss what this mean' (1989; 50).

This quote indicates that, not merely neuroscience, but biology itself is beholden to the notion of consciousness. Introspection is the founding principle of the claim for consciousness which, as with Arnauld and Locke *et al.*, is its justification. Indeed, deference is offered by Dawkins to philosophers who have claimed this justification – without any *scientific* investigation.

\*

What do we have to overcome in ourselves to pass beyond the threshold? We must appreciate the error in supposing the biological function of our apparent experiencing, knowing and being aware is causal. The world is not simply there and we, as subjects, do not have it available to us. It is not that nature sets out to deceive us (*pace* Dennett). It is an error in our interpretation. But in this interpretation we are being the biology we are. The interpretation of

the state *as* consciousness is caused by something we are not, *viz.* conscious. It is a biological cause which has not 'risen' beyond biology as a mysterious evolved property, *pace* Dawkins. The cause is inaccessible to us as (supposed) consciousness.

We have no mental life – no perceptions, no thoughts, no feelings, no conscious ability to act. *This* interpretation will not rid us of being. It will locate what *kind* of being we are in the physical universe. For it is more correct to say: the brain has no perceptions, no thoughts, etc. It is brain functioning we must identify.

The same means of interpretation that 'gave us' the notion of consciousness is going to give us the new model. But one of the great advantages of the new model is that it will demonstrate what interpretation *is* biologically, which consciousness fails to do. Consciousness is a failed interpretation. It is explanatorily inadequate, and biologically impossible.

# The Theory of Brain-Sign

## 5.0 Introducing the theory of brain-sign

This chapter introduces the new model for the brain phenomenon hitherto understood as consciousness, but now defined as brain-sign. Before outlining the model itself, we must set the scene.

### 5.1 Dualism lies not in ourselves but in our theories

To move to the new model we must accomplish two things. These are mutually important and intimately linked. We must develop a new conceptual orientation, i.e. grasp the terms of the new model, and we must completely alter our orientation to ourselves. But this will not be the order in which events occur.

When we learn new things we suppose we acquire new knowledge. This is because we assume we are conscious. But in grasping the new model we see that we have *not* acquired new knowledge. For we now assume that consciousness, the provender of ‘utterly obscure’ knowledge (§4.1.2), does not exist. As the new model proposes, because we are physical, we will acquire a new physical orientation to the world. As a *result* of that, new conceptual structures will develop in us: these are what we have hitherto taken to be knowledge. These conceptual structures do not allow us, or cause us, to act as we have supposed of consciousness (model A, §3.3.1). What causes us to act is our new physical orientation to the world. The new conceptual structures have a different biological description and role, as we shall come to.

We have seen that the notion there are perceptions, thoughts, feelings, and so on, derives from the notion of consciousness *in that* the supposition is we have direct knowledge access to our conscious states which reveal those states *as* perceptions, thoughts, feelings, etc. But if we are not consciousness, then the brain phenomenon ‘we are’ does not provide knowledge for a subject. The conscious subject we have taken ourselves to be has no access to knowledge (for there is none), nor is capable of acting causally upon what we appear to know (for we do not know anything). So, for example, we do not approach the tree because we see it, like it or wish to touch it. It *seems* we appear as the state of seeing, liking and

wishing to touch the tree, but, with the new model, our state in fact neither self-reveals as consciousness (under those descriptions), nor as the causal condition of any action by us as organism.

Clearly there is an ambiguity here. In saying it seems we like the tree, but the liking is not conscious or causal, we appear to be making a basic error. For by definition, we might say, liking can only exist if conscious; therefore to use the term 'liking' and then deny it is conscious and causal is to commit the kind of error we claimed for neuroscience when it uses such mentalist terms for how the brain operates (§4.2). At the moment, until we move further into the explanation, we must stand accused of such errors, for we have not developed the resources to undo them.

The point is, we are separating *functionally* the states of the brain phenomenon which we have taken as our experience, and those states of the brain which cause our actions. The linkage between these two, as we shall describe, is not that they are different aspects of the same state, *viz.* consciousness. So in one move we can dissolve the mind-body problem, and the question of what causes us to act.

But there is a deeper sense in which we must undo the dualism that stalks all notions of mind. The literature is filled with attempts to overcome this dualism, be it topic-neutrality,<sup>70</sup> ontologically unspecific functionalism,<sup>71</sup> identity *à la* Churchland §4.3, or whatever. The (almost, one must suppose) universal supposition of the authors of these accounts is that there *is* a dualism problem to solve. But the dualism problem arises in the authors' own self-interpretation.

If we can convince someone of our theory on some topic, we assume, not that we are a brain attempting to alter another brain (a physical monism), but that we, a person, are attempting to convince another person of what we take to be the case (a mentalist dualism). We assume, in common parlance, that we are trying to change their mind. We are attempting to alter what they take as knowledge *in the sense that* the state we are trying to alter is their knowledge state, i.e. consciousness. So, for example, if we convince them of our theory, our expectation would be that they could repeat it, and we suppose they experience it in a state of mind as ours (more or less). This was the reversion to mentalism Chomsky initiated from behaviourism (1959). Undoubtedly we would suppose their brain had changed in the process,

---

<sup>70</sup> Smart (1959)

<sup>71</sup> As, for example, the conceptual functionalism of David Lewis.

but we would not suppose that *our* brain, independently of our experience, caused *their* brain to change.

But this working assumption is a dualism that cannot be explained away by a theory of how experience could be physical. In being a theory creator and propagator we are *already* subject to an intrinsic dualism of our existence as we take ourselves to be. The question is: Is that way correct?

It is as if, throughout the literature, there is a blind-spot concerning what the issue about consciousness actually is – which we have identified as the predicament (§4.1.2). It does not begin with our theorising; it begins with *ourselves* as theorists. Because we do not see that, we remain dualists in the Cartesian mould.<sup>72</sup> Descartes' dualism was not simply that he considered the mind spiritual as opposed to physical: it was that he proposed there *is* a mind composed of perceptions, thoughts, feelings, etc., by contrast with the functioning of the brain itself.

If there is no knowledge for a subject, then we must look to a different explanation for our (so-called) experience. And as theoreticians (which we all are, even if not scientists), we must work out what theories *are*, and what is involved in convincing others of ours. For holding a theory does *not* entail a causal property which is our experience of the theory (e.g. knowing  $E=mc^2$ ); nor do we convince someone else of our theory by changing their mind so their causal mental state is  $E=mc^2$ . In other words, if we do *not* start from our own dualist assumption about how *we* exist, then we will not suppose there is a dualism problem of minds to be solved.

## 5.2 A brief encounter with history

The new model will describe what function the brain phenomenon actually performs, and what links the brain phenomenon to the brain's causal functioning. But before we begin that

---

<sup>72</sup> Brie Gertler states: 'Appeals to ignorance must be carefully restricted. In particular, objections from ignorance cannot claim that we have no clear idea about qualia, or of the physical, or of the relation between the two; otherwise the claim that qualia are (or are not) reducible to the physical is empty' (2006; 206).

Gertler patently assumes a frame of discourse in which what he wishes to say can be said. If the foundation of what he wishes to say assumes knowledge, then the inverse of it, i.e. no knowledge, will be an empty discourse. But this is simply self-justifying, i.e. that *biological* existence (our existence) *requires* that (psychological *viz.* epistemic) knowledge exists for such discourse. Another discourse, in which our biological states do not demand knowledge (i.e. involve no psychology), is also plausible and, as we shall see, eliminates the kind of problems to which Gertler addresses himself.

description, we must review the work of some prior writers. Although this might seem a diversion, it will demonstrate attempts to move beyond the conceptual shackles of Cartesian mentalism (not mere mind-body dualism) which we have just outlined. These attempts are instructive in themselves, and will be referred to subsequently.

### 5.2.1 Pragmatism

We have encountered the term already. Churchland has given a central tenet as: ‘Crudely, a true proposition is *one that works*’ (§4.3).

As developed by the three founding individuals of pragmatism (though with divergent positions), Peirce, James and Dewey, the idea is that the epistemological problems developing from Descartes, i.e. the search for truth or knowledge in an absolute or apodictic sense, should be replaced by the pragmatic notion that what mental states give us is a means of navigating the world satisfactorily in a behavioural sense. Indeed, the better our mental states are, i.e. the more adequately they are adapted to the world, the more successful our navigation will be. Evolution, and the relevance of science and scientific method generally, are paramount.

Charles Sanders Peirce defined the impact of conception with particular precision.

‘A *conception*, that is, the rational purport of a word or other expression, lies exclusively in its conceivable bearing upon the conduct of life; so that, since obviously nothing that might not result from experiment can have any direct bearing upon conduct, if one can define accurately all the conceivable experimental phenomena which the affirmation or denial of a concept could imply, one will have therein a complete definition of the concept, and *there is absolutely nothing more in it*’ (1931-66; 5.412).

Thus a concept can have a complete definition only by infinite (indefinite is probably a better word) enquiry, experimental enquiry, and only then could it be totally capable of supporting adequate conduct. Peirce is proposing a fallibilist position which demonstrates in finite time our inability to have absolute knowledge upon which we can act, and yet for which concepts (or beliefs) *per se* exist.

Thus it is better for us that we have the concept the earth is spherical not flat, and that the planetary system is heliocentric rather than geocentric. These concepts arise by experience and/or experiment and reasoning, by contrast with mere assumption, choice or authority. But there is no underlying or behind-the-scenes truth about these to which (as human beings) our concepts can aspire.

The position is congenial to a neural network account because a network learns (is trained), but does so progressively. In finite time it is likely that a network could only learn to a degree, and could not learn 'infinitely'. (For what would 'infinitely' mean for a network?)

But Peirce's position raises foundational questions concerning consciousness. Whilst he de-absolutises consciousness in terms of knowledge, i.e. the ('utterly obscure') representation of what *is* to a subject, he does not question its existence *per se*. He does not see, therefore, the irony of his de-absolutisation. For example, Peirce specifically denies there is any such thing as introspection (or reflection), which for other writers is the founding principle of consciousness. He says:

'We have no power of introspection, but all knowledge of the internal world is derived by hypothetical reasoning from our knowledge of external facts' (Haack: 2006; 72).

Yet he does not deny the existence of consciousness.

'Whenever we think, we have present to the consciousness some feeling, image, conception, or other representation, which serves as a sign. But it follows from our own existence (which is proved by ignorance and error) that everything which is present to us is a phenomenal manifestation of ourselves.... When we think, then, we ourselves, as we are at that moment, appear as a sign' (82-83).

In fact, Peirce refers to this as a thought-sign, and later more generally, a *brain-sign*.<sup>73</sup>

But there is unresolved ambiguity here. Peirce wishes 'feeling, image, conception...other representation' to be signs 'present to the consciousness', but does not identify what 'the consciousness' (which is ours) *is* to which representations are present. He struggles with a dualism from which he is trying to free himself. This struggle emanates in these words.

---

<sup>73</sup> Though he also refers to it as a *mind-sign* and *man-sign*, thus exposing the ontological confusion.

‘Word or sign which man uses *is* the man himself. For, as the fact that every thought is a sign, taken in conjunction with the fact that life is a train of thought, proves that man is a sign; so, that every thought is an *external* sign, proves that man is an external sign. That is to say, the man and the external sign are identical.... Thus my language is the sum total of myself; for the man is the thought’ (103).

This astonishing statement, published in 1868,<sup>74</sup> speaks much of the modern era, from Heidegger (early in expressiveness, later with ‘language speaks’ rather than man) and Wittgenstein (the (non-)private language argument), to Derrida (the trace, and Iterability) and Rorty (philosophy as conversation).

Yet it remains ill-defined, indeed into the modern era, for two reasons. Firstly what consciousness *is*, and what its relation to the brain *is*, is not established. Secondly, the *sign* of which Peirce speaks (which is language or thought) cannot be disassociated from the subject’s experiencing *as* consciousness (*as* language or thought). Thus there is no specifically biological formulation, for it remains centred in the (non-biological) mentalism it appears to be trying to throw off.

Moreover, there seems to be a conflict within Peirce’s overall discourse, since he has claimed of conception that it is rational and influential upon action. But a *sign*, in itself, cannot be classified as rational or irrational, and if it is ‘merely’ an external expression, its causal relation to the conduct of the organism is tenuous.

Although Peirce achieves startling vision in addressing the issue, his resources are inadequate to carry it through, and his terms confused.

We should add that Peirce, by contrast with James and Dewey, espoused a scientific metaphysics. James and Dewey effectively disavowed metaphysics.<sup>75</sup> (As such, he would not approve Rorty’s neo-pragmatist position.) Peirce states it thus:

‘What do you mean by there being such a thing as Truth?... You mean that there is something that is SO, no matter if there is an overwhelming vote against it’ (1931-66; 2.135).

---

<sup>74</sup> The referenced article is titled ‘Some Consequences of Four Incapacities’, and was published in the *Journal of Speculative Philosophy* in 1868.

<sup>75</sup> This has been seen as a topic of some controversy, into which we will not probe. E.g. Rosenthal and Bourgeois (1980; 56 and below).

On the other hand, the reach of the old *a priori* metaphysics was constrained by this statement:

‘Any truth more perfect than [the] destined conclusion [of inquiry], any reality more absolute than what is thought in it, is a fiction of metaphysics’ (8.13).

In other words, old *a priori* metaphysics claimed for itself access to a reality behind-the-scenes which is impossible for a mere organism, and fictional as such.

\*

One of the founders of social psychology, also in the pragmatic mould, was George Herbert Mead. Mead’s (1934) position is that:

‘Contrary to Darwin...we find no evidence for the prior existence of consciousness as something that brings about behavior on the part of one organism that is the sort as to call forth an adjustive response on the part of another organism, without itself being dependent on such behavior. We are forced to conclude that consciousness is an emergent from such behavior; that so far from being a condition of the social act, the social act is the precondition of it.... The social act, in its more elementary stages or forms, is possible without, or apart from, some form of consciousness’ (Haack: 2006; 469-470).<sup>76</sup>

The more ‘elementary social stages or forms’ are those of expressive acts, which Mead gives in example: ‘Dogs approaching each other in hostile attitude carry on such language of gestures. They walk around each other, growling and snapping, and waiting for the opportunity to attack’ (466). Although Mead’s emergent consciousness from the social context is no more scientifically classifiable (and therefore comprehensible) than Descartes’ and subsequent individualist, or solipsistic accounts, it does have one advantage. It posits a plausible *explanation* for the existence of consciousness beyond the mere notion of knowledge.<sup>77</sup> But in the internalisation of the mind, Mead has withdrawn from the attempted expressive monism of Peirce, and he builds upon foundationless suppositions.

‘The self which consciously stands over against other selves thus becomes an object, an other to himself, through the very fact that he hears himself talk, and replies. The

---

<sup>76</sup> Also Mead (1934).

<sup>77</sup> Recent followers of this general position are Nicholas Humphrey, and more broadly Michael Tomasello, and indeed the psychologist/neuroscientist Chris Frith, who we have already encountered.

mechanism of introspection' [the plausibility of which Peirce denied] 'is therefore given in the social attitude which man necessarily assumes toward himself, and the mechanism of thought, insofar as thought uses symbols which are used in social intercourse, is but an inner conversation' (481).

We see that not only are talking and hearing not explained (physically-ontologically), even thought is taken to be *using* symbols, rather than *being* symbols. This is quite different from Peirce's 'every thought is an external sign' or 'man is an external sign'. But then Mead requires the self, or I or me, to be a social *construct*, rather than (*à la* Kant) the foundation of the possibility of consciousness. He might seem somewhat Kantian in stating that the I 'is not given directly in experience' (475) but on the other hand he states that 'the "I" is the response of the organism to the attitudes of the others' (*ibid.*) which is not a Kantian consciousness at all. Thus Mead assumes the mechanisms of consciousness without accepting the foundational mechanics of consciousness. It is an unjustified account suiting his social theory. On the other hand, it is an attempt to find relevance for the brain phenomenon beyond solipsistic knowledge.

### 5.2.2 Phenomenology

We must add, however, a note on phenomenology. Phenomenology is the creation of Edmund Husserl. Indeed pragmatism and phenomenology are linked, historically. Husserl read James sympathetically in the 1890s.

There are two general reasons to appreciate Husserl's work. The first is that he explored the notion of consciousness beyond the mere categories of our supposed mental life, thus illustrating what consciousness might entail. The second is that he proposed various reductions, thus to found significance beyond surface (supposed) characteristics. Both of these have been highly influential.

Concerning the first reason, Husserl stated that our perception of a house is not characterised by a two-dimensional appearance – as if we are a mere camera. Constituted as a succession of perceptions, the house is a totality with a front, rear and sides all of which cannot be directly visible in *this* perception (or 'adumbration'), but which we assume, or take into account. Husserl referred to this as the *inner horizon*. Thus the task of perception is infinite (cf. Peirce), since we could never access the entirety of what could be seen.

Concerning our second reason, interest lies in Husserl's various notions of reduction. We will discuss (principally) two here. These are not reductions of consciousness to the brain or the physical world, but reductions within, or to the structures of consciousness itself. (Though Husserl certainly regarded consciousness, or certain attitudes to consciousness, as of the world.) Husserl proposed these reveal both the nature of intentionality and the ego.

He referred to the way humans normally assume the world is there and that they have a psychic life as the *natural attitude*. In this mode, pre-reductively as it were, we take it that our experience delivers the world and our own (e.g.) feeling states as a given, as natural. We do not consider the *work* of consciousness or our ego in delivering this experience; we exist *in the world* innocent of the extraordinary phenomenon which consciousness is. Science assumes this natural attitude, and the findings of science take for granted the natural attitude as an effective way to proceed.

Husserl proposes that to overcome the natural attitude, we perform the *phenomenological reduction*, or *epoché*, in which we reflect upon the dual nature of psychic life. Our conscious states are both directed toward the world, presenting the contents of the world (and our own sensations, feelings, etc.), yet at the same time are *states* of consciousness in and of themselves. This bracketing of the presentation of the world reveals the states of consciousness themselves, and opens them up for reflective analysis. Husserl says, for example:

'Every perception of something immanent necessarily guarantees the existence of its object. If reflective seizing-upon is directed to a mental process of mine, I have seized upon something absolute itself, the factual being of which is essentially incapable of being negated' (1982; 100).

This overcomes Brentano's problem of how there can be 'perceptions' of things that do not exist. My mental life is always an immanence but, being directed, I can be in a state of actual (transcendent) perception (i.e. of the world), *or* (immanent) imagination or hallucination. Consciousness constitutes according to the *kind* of directedness of the mental act.

The distinction between the natural attitude and the phenomenological attitude can be described in Husserl's own words. In the former, our perception of the tree takes it to be what

is out there. But in phenomenological reduction, still accepting that we can see ‘what is out there’, we see also (by bracketing the ‘what is out there’) that the tree is the tree *as perceived*.

‘The *tree simpliciter*, the physical thing belonging to Nature, is nothing less than this *perceived tree as perceived* which, as perceptual sense, inseparably belongs to the perception. The tree simpliciter can burn up, be resolved into its chemical elements, etc. But the sense – the sense of *this* perception, something belonging necessarily to its essence – cannot burn up; it has no chemical elements, no forces, no real properties’ (1982; 216).

Husserl is making a distinction we grasp *via* the phenomenological reduction. Whereas in the natural attitude (in which we normally function) what we see and experience does not have meaning, is simply taken as what *is* as if we were directly in the world *via* our seeing and experience, in the revealed underlying structures of the region of *phenomena*, we appreciate how meaning comes to be. The tree as seen is simply the tree as is; to be meaningful, to be psychologically effective as phenomenal, we recognise the mental act as the *process* of seeing, and what is *constituted* as the seen.

We can extend this explanation by noting that in the natural attitude we suppose that, for example, the tree we see actually exists towards which we could take various postures, for example liking or disliking. But the phenomenological reduction reveals another status. There is no outside toward which we can take different postures. ‘The outside’ is constituted entirely within, and the tree towards which we take the posture of liking is not available for us to dislike; in the latter case our mental state, which is now different towards the tree (a dislike), includes the tree, but the tree as differently constituted.

Clearly, therefore, what is revealed by the reduction does not confuse the world and the mental. Husserl claimed Descartes (and his legacy) had fallen into a fundamental confusion by taking himself to be a thinking *substance* (or thing). Consciousness is not a thing, but a meaning giving region, indeed a truth giving region under *eidetic reduction*, i.e. as (ideal) essences.

Husserl’s idealism does not deny the reality of the world. Indeed, Husserl claimed he had overcome a juxtaposed realism and idealism with what he termed *transcendental idealism*. He regarded himself as having created a new ground for what philosophy could be as an apodictic and truth-bearing philosophical *science*, or *scientific* philosophy. For he had

discovered the point from which philosophy could begin, *viz.* that of the phenomenal region. It is clear that, taking Husserl's position, phenomenology precedes science (at least explanatorily) because the terms in which science can operate are predetermined by the phenomenal region (of which science is normally unaware).

However, Husserl moved beyond the phenomenological reduction to what he (later clearly) termed the *transcendental reduction*. With this he lost many adherents. This reduction was to the *transcendental Ego*. In (roughly) the same way that the phenomenological reduction opened the realm of the phenomena, the transcendental reduction was to open up the realm of the transcendental Ego (beyond, or split from the *empirical ego*) to which the founding of all phenomena could be ascribed.

'If I keep purely what comes into view – for me, the one who is meditating [i.e. reflecting] – by virtue of my free epoché with respect to the being of the experienced world, the momentous fact is that I, with my life, remain untouched in my existential status, regardless of whether or not the world exists and regardless of what my eventual decision concerning its being or non-being might be. This Ego, with this Ego-life, who necessarily remains for me by virtue of such epoché, is not a piece of the world; and if he says, "I exist, *ego cogito*", that no longer signifies, "I, this man exist"' (1960; 25).

But (*pace* Husserl) there cannot be anything that is not of the world. That is precisely the problem. Thus we arrive at a (perhaps inevitable) position as occurred with colours (§3.2.1). For Husserl, what is apodictic (and can ground all knowledge in principle) is the transcendental Ego, and its transcendental experience; but since this secure apodicticity is not of the world, it *cannot* be secure. Indeed, since for Husserl the Ego is self-constituting, we have no point from which to determine whether such an Ego actually exists which *could* ground knowledge. As with the sole proof of consciousness being our awareness of being conscious, Husserl's self-constituting transcendental Ego is self-endorsingly tautology.

Thus although Husserl specifically denies that what we (e.g.) see are representations, the *predicament* (which we said related to the problematic condition of knowledge as representation of what *is* to a subject) also functions for Husserl's transcendental Ego. We cannot accept anything consequential outside the world, or indescribable as part of the world, particularly that which founds our (supposed) existence as knowers (though not representational knowers). Somewhere the argument has gone wrong. As Quentin Lauer has

said: 'Phenomenology began as a quest for objectivity, but what it has done is to define objectivity in terms of what it has found' (1965; 65). Tautology again.

That wrong argument discredits consciousness (as the province of knowledge) entirely.

\*

Our last subject was Husserl's sometime assistant and colleague, Martin Heidegger. Heidegger's underlying concern was different from Husserl's, and marked a profound change of emphasis in phenomenology, and subsequent philosophy. Whereas Husserl was fixed upon the subject and its route to knowledge, Heidegger was in pursuit of being. Both were concerned with an absolute, but Heidegger's was the route to, and manifestation of what *is*.

According to his own version of events, Heidegger changed his approach at the turn (or *Kehre*), roughly the mid-1930s. Prior to that, and in the two sections that were published of his 1927 *magnum opus*, *Being and Time*, he outlined what he termed a fundamental ontology of the being of man in his/her being, or as he termed the entity that had such a being, *Dasein* (literally, 'there is', or 'being the there').

One striking aspect of this text is its apparent pragmatic orientation. There has been a supposition that he was influenced by Theodore Lask, who had developed a pragmatic approach on reading American pragmatist texts (see Theodore Kiesel, 2002).<sup>78</sup> Rüdiger Safranski, in his biography of Heidegger (1998), notes pragmatism as one of the key philosophical positions of the time.

But the pragmatist comparison must be treated with caution. For Heidegger's discriminations are distinct from pragmatism's recoil from metaphysics. Heidegger's position was that it was precisely the history from Plato onwards in Western conceptualisation that ends in the wayward orientation of pragmatism.

Heidegger was both indebted to Husserl for his ontology, and opposed to Husserl's own approach to phenomenology in *Ideas*. *Being and Time* is at once a homage to Husserl (recognised in its dedication), and a critical commentary.

---

<sup>78</sup> The paper by Kiesel, author of a monumental work of scholarship, *The Genesis of Heidegger's Being and Time*, was first published in 1993. In *The Genesis*, Kiesel states: 'Heinrich Rickert, teacher of both Lask and Heidegger, observes in his final report on the habilitation work – he was the director, the *Doktorvater* – that Heidegger "is in particular very much...obligated to Lask's writings for his philosophical orientation as well as his terminology, perhaps more than he is conscious of"' (1993; 25).

We have seen, with Husserl, that the phenomenological reduction reveals a way of grasping consciousness which the natural attitude conceals. Heidegger, as it were, reverses Husserl's formulation, but this results in no less a fundamental reconstruction of human self-conception. The natural attitude is seen, not as a naïve way of existing, but *the very nature* of existing, now termed *facticity*. It is facticity that must be uncovered and elucidated. This uncovering focuses upon the nature of our actual practices, our *Being-in-the-world*, as Heidegger speaks of it. And it is in this sense that the apparent association with pragmatism appears. Heidegger says:

'The "factuality" of the fact of one's own Dasein is at bottom quite different ontologically from the factual occurrence of some kind of mineral, for example. Whenever Dasein is, it is as a Fact; and the factuality of such a Fact is what we shall call Dasein's "*facticity*".... The concept of "facticity" implies that an entity "within-the-world" has Being-in-the-world in such a way that it can understand itself as bound up with its "destiny" with the Being of those entities which it encounters within its own world' (1962; 82).

Husserl had rejected representationalism. It can be seen, from this passage, that Heidegger further rejects knowledge. Knowledge is not something a subject gains as a separate entity from the world for whom (in some mysterious way) the world becomes apparent *via* consciousness. The Being of this being, Dasein, is such that it is already in the world, and as such its understanding of the world already is (in some way). Intentionality, for Heidegger, is not the directedness of one's mental states on some object (as with consciousness for Husserl); in that sense there are no mental states, or minds as which they occur. Intentionality, as Heidegger conceives the concept (the word is used in *Being and Time* only when discussion of Husserl and Scheler occurs)<sup>79</sup> is to be understood in terms of how Dasein *is* in the world, in that in relating subject and object, the subject's being is constitutively established. Husserl's transcendental Ego has no place here.

Heidegger's reference, in the passage, to the mineral as which Dasein is not, is not arbitrary. As with Husserl, but with another emphasis, Heidegger accuses Descartes, and the subsequent tradition, of confusing the nature of Dasein as an alternative world of the mind conceived as an equivalent state or condition to objects in the world with their properties. The entire rhetoric of mind stuff (e.g. perception is taken as a kind of mental *property* like the

---

<sup>79</sup> This point is made by Frederick Olafson (1987).

hardness of a mineral) has obliterated the nature of human being, and it begins with the later Greeks. Because our language is so tied up with the thingness of worldly things, we have no adequate vocabulary to express what Dasein is. Thus Heidegger has to construct a terminology adequate to express Dasein's existential states, and existence generally. (As Heidegger would say, Dasein exists but a mineral does not, thus reserving the term *existence* for the kind of being that has existence as its way of being.)

Part of Heidegger's account of Dasein appears to bring it even closer to pragmatism. This is the way in which Dasein has the world and its objects revealed through its engagement with the world in a practical way. But it is precisely here that a clear differentiation can be made. He says:

'If Dasein is ontically constituted by Being-in-the-world, and if an understanding of the Being of its Self belongs essentially to its Being, no matter how indefinite that understanding may be, then does not Dasein have an understanding of the world – a pre-ontological understanding, which indeed can and does get along without explicit ontological insight?' (1962; 102). And: 'When we concern ourselves with something, the entities which are most closely ready-to-hand may be met as something unusable, not properly adapted for the use we have decided upon. The tool turns out to be damaged, or the material unsuitable. In each case equipment is here, ready-to-hand. We discover its unusability, however, not by looking at it and establishing its properties, but rather by the circumspection of the dealings in which we use it' (*ibid.*).

This seems close to the pragmatists' conception, that our knowledge about the world is rooted in practical action, and what we believe about the world is an aid to our practical conduct, particularly if it is successful. (Indeed, Heidegger is referring to tools as *pragmata*, the Greek for things in a usage situation, rather than *res*.)

But Heidegger's is a different orientation. For it is not a (pragmatist) question of whether our beliefs are useful in action: it is how the world and its objects are revealed *in* action. Heidegger does not say that we cannot know about the world, but rather that knowledge is revealed initially not as formally construable, for example, as having properties. It is revealed by our engagement in the world, of which practical action is an example. It does not depend upon consciousness as a transcendent state beyond the world revealing it to a 'worldless subject' (which may not need a world at all, as Husserl had said). In other words, the word *knowledge* means something other than a condition of consciousness. It is, for

example, neither subjective nor objective, for there is no subject for whom it exists set off against the world as objective.

Of the numerous topics we could explore, one is crucial to our concern. In the history of mind, there has been a fundamental difficulty: *other* minds. From the Cartesian beginnings the problem is that we are aware of ourselves from an internal perspective, but how can we be aware of others in an equivalent way (§2.4)? Indeed, if the world is given to *us*, how can it be given to others equivalently without there being multiple worlds? Husserl struggled with this without resolving it, though he was adamant that others coexisted.

Heidegger's position somewhat alleviates the problem because he does not posit a subject (of the mind) to which the world is given, but reverses the situation such that we (as Dasein) are already in the world. As such, we can be with others who are also in the world.

'The expression "Dasein"...shows plainly that "in the first instance" this entity is unrelated to Others, and that of course it can still be "with" Others afterwards. Yet one must not fail to notice that we use the term "Dasein-with" to designate that Being for which the Others who are are freed within-the-world. This Dasein-with of the Others is disclosed within-the-world for a Dasein, and so too for those who are Daseins with us, only because Dasein in itself is essentially Being-with' (1962; 156).

Although Heidegger refers to the situation of Dasein in the world as essentially mine, he is not solipsistic about Dasein's relation to others. And in that regard, others take it that we exist in a similar manner.

Although for many Heidegger's account is congenial in describing our existence in the world, and addresses problems that the conceiving of ourselves as (dualist) minds had posed, it does not address the nature of Dasein as emerging from matter, i.e. the brain. Indeed, it establishes an extraordinarily awkward position in that Heidegger explicitly demarcates Dasein as contrary to matter *by* existing. If we take the universe as entirely physical, then we have the contradiction that existence emerges from non-existence. Moreover, Heidegger depends fundamentally upon the notion of illumination for how the world is given.

'When we talk in an ontically figurative way of the *lumen naturale* in man, we have in mind nothing other than the existential-ontological structure of this entity, that it is in such a way as to be its "there". To say it is "illuminated" means that *as* Being-in-the-

world it is cleared in itself, not through any other entity, but in such a way that it is itself the clearing. Only for an entity which is existentially cleared in this way does that which is present-at-hand become accessible in the light or hidden in the dark.... *Dasein is its disclosedness*' (1962; 171).

Heidegger wishes to dispense with mind terms because he wishes to dispense with mind. A way of doing this, for the term perception and indeed knowledge, is to introduce the notion of clearing and disclosedness which *Dasein* is. The need for a new vocabulary is understandable, given the revolution Heidegger intends. But...

Again we have the unresolvable problem of how the brain can glow such that what is given to (or in) the brain of man as the there. Of course, if we refuse to accept scientific priority, that prior to any possible science, what is disclosed as *Dasein* is what *is*, our problem goes away. Indeed, Heidegger is clear on this point.

'The sciences still speak about the Being of beings in the unavoidable supposition of their regional categories. They just don't say so. They can deny their origin from philosophy, but never dispense with it. For in the scientific attitude of the sciences, the document of their birth from philosophy still speaks' (1969; 377).

But this might be seen as no more than the very claim to power, in this case of philosophy over science, that Heidegger has warned us about, and proposed has led us astray since Plato. Moreover, this particular claim to power appears self-contradictory, for philosophy, as Heidegger conceives it (and characteristically of philosophers, his claims are self-justifying), cannot account for how the existent can arise out of the non-existent. It is precisely science which renders Heidegger's ambition, revolutionary though it is, as void *in its own terms*.

### 5.2.3 Summary

Each of these theorists attempted to destroy the Cartesian notion of mind. Despite their lack of success, they certainly did not leave the notion of mind untouched. The problem was, they were not revolutionary enough. We can see how difficult the task is, to escape the cultural suppositions in which we are encased. Their efforts were hampered by remaining in the solipsistic tradition which is consciousness in essence. I.e. the world is present to *me*: I,

therefore, am (the centre of) the world (cf. Descartes in Chapter 1, & §2.2f Heidegger on Kant). This was so even though one of their aims was to account for other beings beyond the subject of consciousness. The only one proposing a possible reason for the existence of consciousness (with a somewhat biological conception) was Mead in terms of social function. Nevertheless he did not escape the mentalism of the individual.

To identify a solution which avoids our cultural fixation, we must escape mind and solipsism entirely. We must see what problem biology *qua* physicality has to solve.

## 5.3 The scientific model of brain-sign

### What biological problem is to be solved in the brain?

#### The super-organism

If there is no consciousness, as we conjecture, then obviously the powers of consciousness do not exist. Therefore an account of organisms must not presuppose what consciousness has been deemed to make possible. Only then can a scientific theory of the brain phenomenon be developed. The question is, therefore: What function does the brain phenomenon perform?

#### 5.3.1 Preliminary orientation – the natural and scientific dispositions

Clearly this question is not the first in the hierarchy of questions. For as we have indicated, we must first ask what it is to ask questions, or more precisely: *What* asks questions? A conscious person asks questions because they are conscious. But if there are no conscious persons, then questions cannot be asked by them. That there appear to be questions asked invites us to consider how *we* stand in relation to question asking. What are *we* that seem to ask questions?

It might be said that questions are asked by the *unconscious*. Perhaps this is what Allen and Reber, or Edelman and Tononi might propose (§2.7). After all, since we do not know the origin of what causes us as consciousness (§4.1.2), it might be supposed that the cause of us asking questions is really the unconscious without our realising. But this proposal is evidently wrong if there is no consciousness. The only available thing that could ask questions is the neural brain. But this answer would be absurd. The neural brain cannot ask questions: it is a mere physical object. The only possible conclusion is, therefore, that there are no questions. Which means there is only the appearance of questions.

But then the question must be asked: To whom can there be an appearance of asking questions? And again we are forced into an impossible situation, for the only entity for whom the appearance of questions could exist is the conscious person. And since such does not exist, there cannot even be the appearance of asking questions.

But is this so surprising – to find we have no point of departure in asking the most straightforward of questions (about asking questions), which itself takes for granted we are talking about a function of consciousness, when there is no such thing? Since we have shown we have no means to verify the brain phenomenon we are is knowledge (§§4.1.x), whatever the phenomenon is cannot be such as to sustain the kind of processes we have supposed for that which has knowledge (which questioning might reveal), *viz.* the conscious person. This is entirely in accordance with the proposal that there is no perception or thought. To get to what actually happens in the brain we must take a different route. Both Husserl and Heidegger set up new routes, and so must we, though not the same routes, for our assumptions are quite different.

Yet the contributions of Husserl and Heidegger are not irrelevant. Since we appear to ask questions, what we are, as a brain product includes the appearance of asking questions, even though it is not even an appearance. What this means (since I cannot get before or behind my state to interpret it directly) is that I am a brain product for its (the brain's) purposes. And the only way of grasping what that product is (although I cannot grasp it in fact) is to establish a scientific model which accounts for what appear to be the facts of the situation. In that model, what I am (which is not consciousness) will become apparent (though not to me as conscious).

Since I am a phenomenon of the brain which is not consciousness, yet which has been interpreted as such with associated features and knowledge, we shall follow Husserl in naming it: *the natural disposition*.<sup>80</sup> It is not an *attitude* towards the world and itself, for it is a phenomenon of its creator (the brain), not, as it were, a self-generation as a self (as Husserl supposed). In model terms we will simply suppose the brain phenomenon is built in such a way as to *dispose* the organism with certain features – and this for a biological function as yet to be established.

We shall distinguish the natural disposition from the *scientific disposition*. In the latter case the nature of the brain phenomenon has been uncovered as a physical phenomenon, subject to lawful interpretation in a scientific manner. The first posit of the scientific disposition is that the brain phenomenon is not consciousness and does not function on the assumptions associated with the natural disposition. The second posit is that, *as* the scientific disposition, what occurs *as us* is not consciousness. Thus in positing that we are not

---

<sup>80</sup> In fact, Kant used the term in the Introduction to the first *Critique*.

conscious, we must suppose that our state has to be reinterpreted biologically from that of the natural disposition. Crucially, we are not about to claim knowledge of ourselves (because we are not conscious) – and this will eventually reflect upon the nature of science itself.

We will follow Heidegger also, in that we must suppose that *as is*, the natural disposition needs to be interpreted. But this interpretation is not going to give a fundamental ontology. It will invoke a reduction to the physical universe as whatever the brain phenomenon is in biological terms. In achieving this, the scientific disposition will be elaborated.

What is implied here is that in performing the interpretation and reduction, we are acting in accordance with being that self-same brain phenomenon, and cannot escape to some neutral realm of exploration of ourselves as if we (as natural disposition) were an object to be studied, as it were, scientifically (the kind of mistaken supposition involved in psychology).<sup>81</sup> Again, this is because we cannot claim we are a special type in the universe which is existent (*à la* Heidegger), for we can make no claim to raise ourselves beyond the physicality we are. Our model will always be a conjecture from out of the physical universe itself. Thus what we are, as the brain phenomenon, must accord with what the physical universe can be. In taking this approach, we are acting in a way customary in science.

### **5.3.2 The grounds of the biological model**

We shall make a number of fundamental claims. A key claim is that organisms, as an element of their ability to survive and reproduce – the sole biological rationale for their existence (cf. Darwin) – need to communicate with each other. The reason being that, in evolutionary terms, there are many survival advantages in acting as a collective. This is not a novel claim. However, when consciousness is denied, its significance comes sharply into focus.

#### **5.3.2.1 The literature confusion**

We see this already in our reference to Mead. Mead says: ‘Dogs approaching each other in hostile attitude carry on such language of gestures. They walk around each other, growling

---

<sup>81</sup> This kind of critique of psychology was made by Husserl, from his standpoint on the natural attitude.

and snapping, and waiting for the opportunity to attack.’ This behaviour, he proposes, occurs before the need for consciousness.

But as we have seen previously, his statements on the data are not clear – here in one way probably, in the other definitely. For how is dog A to interpret the gesture-language of dog B? If the supposition is that dog A *understands* dog B’s gestures, already there is a requirement for consciousness – unless Mead is claiming that such activity takes place non-understandingly, i.e. as a neural process without consciousness. Indeed, this he seems to be saying. ‘The conversation of gestures is not significant below the human level, because it is not conscious, that is, not *self-conscious*’ (Haack: 2006; 470). This seems to imply that understanding requires self-consciousness (cf. Davidson, §3.1.4). But in referring to growling, with the necessary requirement that growling be *heard*, consciousness is already presupposed. And Mead continues, now in brackets: ‘though [the conversation of gestures] is conscious in the sense of involving feelings or sensations.’ This indicates Mead’s ontological confusion, resulting from his analysis in the natural disposition *which cannot reveal actuality* (as we shall see).

In *our* natural disposition we take the world as does Mead: we hear the dogs growl and we suppose this is causal. But, as with Mead, our mistake is not conforming to the appropriate biological model of ourselves, or the dogs.

In the natural disposition we might regard some animals as self-conscious: some apes, and dolphins. But we cannot make an ontological discrimination in the physical universe between the existence of (conscious) sensations (e.g. hearing) for dogs and being *self-conscious*, as are humans. Surely there may be a gradation between species in whatever the brain phenomenon is (indeed, we shall explain exactly that), but not of such a fundamental kind. For, for a sensation to occur in dogs (as Mead proposes), there must be a conscious subject for whom the sensation exists. Indeed, as we have seen, the foundational justification for consciousness throughout history is a subject able to be conscious of being conscious. As a result there cannot be conscious sensations which cannot be, in principle, self-consciously identified. So either the proposition must be that dogs are self-conscious to have conscious sensations, which is apparently (empirically) false, or that dogs do not have conscious sensations. So Mead’s analysis simply makes no sense – unless he were to bring forward some vastly more fulsome and fundamental ontological analysis.

Mead's confusion concerning animals is evident throughout the literature.<sup>82</sup> This is because there is no adequate ontological-biological ground from which to *begin* an analysis of the brain phenomenon. For example, in modern terms, a writer who ascribes consciousness to language, including the derivative feeling-concept (e.g. sensations), is Euan Macphail. 'Language is an instance of undeniable behavioural complexity...from which it is possible to derive an account of the origins – via self-consciousness – of feeling-consciousness' (1998; 235). As a result of this, Macphail denies that either animals or the new-born are conscious.

On the other hand recent exchanges in *Behavioral and Brain Sciences*, initiated by a paper by Ned Block (2007), are replete with the literature confusion – as some commentators note. Block himself is guilty. For example, he wishes to deny that phenomenal consciousness can only exist if reportable. This, he claims, is because those with such states as profound global aphasia cannot report their consciousness, but this does not mean (he proposes) they are not conscious. He then states, without making any distinction, that: 'Come to think of it, the same applies to animals and pre-linguistic children' (484). Tyler Burge concurs with this, but from a different position. 'How could it be apriori that animals that lack propositional attitudes cannot feel pain?' (2007; 500). Burge, from the natural disposition, proposes the issue must be empirical.

So we see that, as with colours (§3.2.1), the consciousness identified in the natural disposition possesses no generally accepted ontological grounds for existing.

*However*, if one organism is to communicate with another (which surely happens) and consciousness does not exist (i.e. if we by-pass the consciousness confusion), then we must assume that a non-conscious brain, *via* its sensory input, is all that can process communicative material.

### **5.3.2.2 Communication supersedes the individual**

We must go further along this route and consider a founding assumption in the literature of communication. It is generally supposed that the physical differentiation of organisms marks their functional boundary (cf. §2.4). Thus *communication* is deemed to be interaction

---

<sup>82</sup> As we shall come to, as representative examples, neither Marc Hauser's comprehensive *The Evolution of Communication* (1997), nor Maggie Tallerman's (ed.) *Language Origins: Perspectives on Evolution* (2005) have one index entry for the word *consciousness*, yet the contents of both consistently refer to states which presuppose the existence of consciousness.

between two functionally separate units. Now it is clear that two separate organisms are physically separate: but what does communication effect? In both an evolutionary and functional sense, communication overcomes a *lack* in the individual. It renders two organisms *not* separate precisely because communication organises communal behaviour superseding the individual boundary. Communication is the transparent binding of two organisms into one (transparent because neither is 'aware' of the mechanism). In this binding they become the *unit of the communicated*. Although they are not bodily connected, they are connected physically because the process of communication is physical, as we shall see.

The significance of this is paramount. It is not generally expounded in the literature because the cultural notion of consciousness isolates the individual with a mind which is *self*-orientated. And this is true whether or not other creatures are deemed to operate with minds – which, as we have said, is pervasively confused in the literature. When theoreticians observe the interaction of two (or more) organisms, what they *see*, in their natural disposition, is two (or more) organisms.

Here is an example. Of Darwin, Marc Hauser states: 'He argued that expressions were designed to convey information about the signaller's emotional or motivational states, with some signals reflecting an underlying ambiguity or conflict between different emotional states...such as fear and aggression' (1997; 19).

Darwin's approach begins from the mentalist individual rather than from the outcome of the communication process for the physical collective. It does not take into account that, in the physical universe, the process of communication depends upon the physical means of communication which is an evolutionary incorporation. The modification of organism A's behaviour by organism B's signal depends upon both organisms being in a physical universe in which the passage of communication uses electro-magnetic radiation and/or compression waves effecting it. What results from A's signal is a neural modification in B, and thence (in principle) a physical modification of B's behaviour. As we shall see, A and B act as a collective without any question of motivation or emotion, or these being communicated to a knowing recipient *as* information – because consciousness does not exist, as which they are supposed to be sustained. Darwin's explanation, as expressed by Hauser, is not an explanation of physical actuality. It is not, therefore, scientific. It depends upon the posture of the natural disposition.

Hauser continues: ‘Concerning communication, the grandfathers of ethology – Lorenz, Tinbergen, and von Frisch – largely accepted Darwin’s treatment, especially the idea that signals were designed to communicate information about the signaller’s motivational or emotional states’ (19-20).

Thus the approach was continued. But the theory runs into trouble because, during the early days of ethology, ‘research was influenced by a conceptual framework that saw the forces of natural selection operating at the level of the group’ (20). The problem was altruism. Why, for example, would an individual express an alarm call to others in the face of predatory danger, when making that call would expose the individual to greater threat because it had identified itself to the predator?

‘[By] thinking in terms of the individual, or more appropriately, its genes...selection would readily favor a mutation that caused an individual to withhold the alarm call, run for cover, and save its own skin.... This interpretation, with selection operating at the level of the individual or gene, has become the dominant perspective in the study of animal behavior and was largely responsible for changing how biologists think of about communication and design’ (*ibid.*).

We shall not, at this point, pursue the history of the topic because the relevant point has been made. Whether or not there is an evolutionarily valid account for the behaviour of so-called altruism, the literature, by focussing upon the individual or gene within the notion of communication, whether of deterrence or cooperation, has forced a fundamental ambiguity or discontinuity in biology. The mere ability of communication, built into the bio-functional organisation of the individual and genetically inherited, is already a mark of a collective practice which is a necessary *addition* to the explanation of behaviour from the individual or (‘selfish’) gene.

Communication is a genetically inherited capacity superseding the individual behaviourally. It is inappropriate, therefore, to discount this by wholly founding behavioural explanation in the individual or gene. We may of course explore the specific function and mode of any given communication practice, but it will be built upon a fundamental biological characteristic (its physicality) which *itself* must be exposed. An individual’s genes already sustain the inherited propensity for evading individual limits. This can be exposed by disregarding the cultural supposition of consciousness or mind (with its solipsistic orientation), and thereby escaping the explanatory limits of the natural disposition.

### 5.3.2.3 Communicative modalities – the biological differentiation

Although language is usually differentiated from mere growls because of its (e.g.) syntax, complexity or capacity for abstraction, we see (in the Mead discussion) that growls (as hearable) have the same requirement for consciousness as the language we hear. This is not the case for all communicative modalities. Indeed, there is a fundamental differentiation to be made in the modes of communication which we will identify.

### 5.3.2.4 Signification – static and dynamic communication

We identify the mechanism for organismic communication in signs, which are physical, and their placement in a non-conscious structure. Within that structure we are going to make a fundamental claim, which is a novel claim in biology itself (by contrast, of course, with signs, or signals, *per se*).

The differentiation cuts across accounts of organismic communication generally because it posits a new way of breaking down the data. We propose there are two fundamental states of communication between organisms. The first we refer to as *static*, the second *dynamic*.

In the static case, the sign or signal from one organism to another causes a specific (or determined) kind of resulting behaviour. There needs to be no *interpretation* by the recipient of the state of the signer. The sign causes a resulting behaviour of the receiver (in principle, because it might not work in practice, cf. Millikan's notion of *proper function* (1984)).

In the dynamic case, the behaviour of the signer is not necessarily unambiguous. Indeed, the behaviour of the signer is not necessarily a sign, although it must be responded to by the receiver for their joint behaviour to *be* cooperative.

The difference influences the cognitive demand upon the brains and general biology of organisms. This does not mean, however, that static communication is necessarily simple, or that dynamic communication is necessarily complex. The cognitive demand relates to the *nature* of the communication process.

### 5.3.2.5 Signification of affective states and reference in the literature

The issue of the cognitive demand of affective states and reference is of course addressed in the literature, but not as we approach it. The differentiation found there generally is driven by suppositions about human language, taken to be capable of extensive reference. However, though nonhuman communication is taken to be limited, it is not supposed that nonhuman communication only concerns affective states (e.g. emotion and motivation) with *no* reference. Hauser says:

‘Though nonhuman animals clearly convey important information about their affective states, they also appear to convey information about objects and events in the external environment’ (1997; 62).

Our question is: How is communication effected? It is worth commenting again on Hauser’s words. He seems to propose that information is conveyed between organisms as ontologically separate from the environment in which they function. For example, their affective states are regarded as internal, and objects and events (in the environment) are external. While there is a differentiation to be made between organism and non-organism in the communication process, our analysis of behaviour in communication is taken to be *of* the ontology of the physical world from which organisms are *not* to be differentiated. While this might seem obtuse for the natural disposition, from the scientific disposition it will help clarify what actually happens.<sup>83</sup>

### 5.3.2.6 Static communication

Our example of static communication is the bee dance. We use this because, not only is it well known and extensively documented since the pioneering work of von Frisch (1967), it is also plausibly the most complex communication of the static kind, employing, as it does, (supposed) reference to the world. (An interesting discussion of the controversies in the history of the topic can be found in Shettleworth (1998).)

---

<sup>83</sup> Thus Hauser introduces but does not justify ontologically his claim that ‘Studies of child language tell us that sophisticated word use emerges with sophisticated development of other cognitive abilities, including the ability to recognize mental states in others and take advantage of rich conceptual representations.’ (1997; 67). We will come to this.

A principle function of the bee dance is to convey, from the forager bees who have identified the source of nectar, its location to receiver bees so they can set off to collect nectar too. According to Hauser, the method of conveying the information is ‘visual, auditory, olfactory and tactile’ (1997; 497).

It is evident from his language that the supposition is that bees can see, hear, smell and feel. In other words, the supposition is that they are conscious as a result of which each of these sensibilities can function *per se*. But this is surely deeply improbable. Kim Sterelny (2003), reviewing the literature, states that he does not consider it necessary that bees have any kind of internal representation, or mind. Indeed, Hauser himself states that ‘there is no evidence that I am aware of, that communicative signals [of organisms apart from humans] are capable of creating or evoking an image in the minds of listeners, or that signallers have an image in mind when they communicate’ (1997; 509) (cf. §5.2.2.1). Though this passage is aimed beyond bees (and appears to contradict the notions of ‘visual, auditory, olfactory and tactile’), it must apply to bees, and must include *any* kind of seeing, etc.<sup>84</sup>

From this we would agree that all communicative modalities take place in purely physical terms; i.e. the various communicative modalities programme neural states physically, causing an effective response.

The dances are principally a round dance conveying information about nearby sources of nectar, and a waggle dance for more remote sources. But in addition, information is conveyed by both movements in the hive (for example, what is called tremble dancing), wing beating in the air, and sound emission of both foragers and receivers. As Hauser says: ‘These extremely small creatures, with their relatively simple nervous systems, are able to communicate quite explicit information about the distance, direction, and location of food’ (1997; 503-504).

Let us considered just one modality. Hauser refers to studies which have identified that in either the round or waggle dance,

---

<sup>84</sup> But note that research may still assume such states exist. Heidi Lyn refers to apes having minds with mental representation (2007; 468), and this is based upon associations and errors in the use of words by apes on the supposition that apes, like humans, have a mental architecture. This is hardly evidence when the notion that humans have minds is unproven; but of course Hauser supposes humans do have minds. Hauser’s confusion of what constitutes a mind, and to what degree animals have them, remains in his *Scientific American* article of 2009.

‘When foragers return to the hive...they emit a...sound [which] has a fundamental frequency of 200-300Hz, a pulse duration of 20msec, and sound-pressure level of 73dB measured at approximately 1-2 cm from the [dancing] bee. The sounds emitted appear to provide information about direction, distance, and even profitability. In contrast, the sound of the tremble dance consists of a fundamental frequency of 350-450 Hz, an average duration of approximately 140msec, and a sound pressure of 80dB recorded at 1 cm from the dancing bee. The function of these sounds is to arrest dancing in others and increase the rate of food storage’ (502-503).

We are faced with a similar situation as with the mantis shrimp (§3.2.1.1), but now concerning communication. For *sound* to exist, there needs to be a hearer, and bees surely cannot hear. Moreover sound, with its requirement for a hearer, implies some kind of intelligent use (and perhaps more than one) of the *representation* that sound is for the conscious hearer. Thus Hauser, presumably as a normal literature usage, refers to sound because he assumes (without considering it) that we will associate the frequency of compression wave emission with the word ‘sound’, because *we* supposedly hear sound from compression waves (on the supposition we are conscious). But in scientific terms this totally obscures the topic.

We must avoid the obscurity in Hauser’s usage. As a static form of communication we propose that each specific modality of the dancing bee’s performance, in conjunction with its context, causes a specific behaviour modification in the receiver. The physical programming is such that, mistakes aside, there is no (conscious) dance *interpretation* required, because forager and receiver bees are part of the physical nature of the unit of the communicated. In other words, there is neither reference nor representation *for* forager or receiver subject, and even the word ‘communication’ must be treated with caution lest it evokes something beyond pure physicality (e.g. consciously understood information). Seen this way, there is a rigidity of physical causation from forager to receiver *via* the interceding physical strata (compression waves). The unit of the communicated is, of course, remarkably complex for such simple creatures: but it is entirely pre-structured to effect the result.

To be explicit, neither foragers nor receivers *know* what they are doing. They are embedded within the physical world, and act according to their physical being.

### 5.3.2.7 Dynamic communication

But now we must consider whether *any* creature knows what it is doing (though we have denied this, §4.4). Certainly, in terms of communication, this cannot be associated with the distinction between affective states (e.g. anger) and reference. The most complex creature in a behavioural sense is the human with, it is supposed, a highly complex referential medium, *viz.* language. But humans can react directly to an emotional state of anger in another (as the natural disposition has it) as swiftly as any simpler creature. So what is the criterion for dynamic communication if it is not this contrast of these states?

The issue is ambiguity in the behaviour of another. This need not be deliberate, but merely a lack of explicitness – by contrast with the bee dance.

At this point we introduce the notion of the *super-organism*. A super-organism, as the expression implies, involves an organism-type functioning above the level of the individual. The term is applied where individuals cannot survive independently of the collective (as in the ant community), but also where mere social benefit is gained by cooperation. In the latter case the benefit arises from the range of skills, or indeed sheer numbers in the group which improve its survival potential. (The essence of sociology.) The bee community, with static communication, is a super-organism.

*Our* usage of the term concerns the *mechanisms* of communication. We have shown such mechanisms in the bees' static communicative behaviour (the dance); we now show how organisms have evolved for dynamic communication with *other* mechanisms which have been mistaken as consciousness or mind. Indeed, we will show (eventually) that the very notion of consciousness or mind in an individual is part of the communication structure, but imperceptible to us while 'we' function *as* the communication structure, i.e. in the natural disposition. To see this requires us to engage the scientific disposition.

One problem we will eliminate is whether and to what degree animals are conscious. No animals are conscious. However, long before humans in the evolutionary scale, dynamic communication evolved. Thus one of the advantages of our proposal is that we can establish evolutionary continuity with our biological 'predecessors'. The non-human example we choose is hunting lions.

The problem is obvious. Two lions pursue a gazelle. Each must take account of the other's movements to be able to corner the gazelle, and so attack it directly. If we suppose

lions have minds or consciousness (as in the natural disposition), then we resort to the concepts of perception and understanding: each lion sees and understands what the other lion is doing in relation to the gazelle's movements. But there are two problems. The first is that there is no effective account of perception or understanding in the physical universe (i.e. reductively). The second, on the other hand, is that, while we might say that consciousness is irrelevant, that mere neural processes accomplish the task, when we compare the static bee dance, and its behaviour-influencing signifying structure, with the lion hunt – where the movements of the lions cannot be considered equivalently as signs – we must ask how communication between them *can* take place.

What does the super-organism achieve in its unit of the communicated as a structure of mere physical states? The aim (or purpose)<sup>85</sup> of bees (though it is only an *aim* as we interpret it in the natural disposition, cf. Slezak discussion §§3.1.3 & 3.1.4), is to acquire and store nectar. But they operate as one mindless physical structure. Must we hypothesise, by contrast, that lions do understand their common aim of killing the gazelle for survival nourishment? But we have no physical mechanism available for understanding to take place. Do they, therefore, merely understand the *hunt* process: that the movement of lion A to the left at this moment means that lion B must move to the right in relation to the gazelle? But we do not know in what terms to specify such an understanding, in this case because lion A has not *signified* an intention in moving to the left that lion B should move right. Even so lion B appears (to us in the natural disposition) to take lion A's move as an intended action in the hunt relative to the gazelle and the other lion *as if* it were a sign causing lion B's response. Are we then to say that lion B has a cognitively valid non-conscious (i.e. neural) grasp of A's move? But again, since no sign is involved, no non-conscious (or purely neural cognition) can take place.

Let us summarise the situation.

1) We seem forced to resort to perception and understanding in the lions because we have no other model to explain how they can cooperate successfully. We take the lions to be like ourselves, and we suppose we perceive and understand, and that is how we get about in the world, cooperating with others. The natural disposition in action. But if consciousness is language dependent (cf. Macphail), for only thus does a subject exist who could be consciousness, then lions have no language. And if they are conscious without language or

---

<sup>85</sup> Recall the quotation from the science orientated Robert Crease at §4.1.

being a subject, what does that mean (we would ask Block and Burge, §5.2.2.1)? And to implant consciousness directly upon the neural (the *personal level* to the *subpersonal level* as does Susan Hurley (2008) after e.g. Dennett (1969; 1991), §4.2) without an explanation of consciousness, is otiose.<sup>86</sup>

2) Then again, what benefit would there be in making cognitively valid decisions beyond purely neural processes by (conscious) images of the gazelle and the other lion in each lion's head? This could only be explained if we had an engineering account of how it would work. Such an account is not available.<sup>87</sup> (And it risks epiphenomenalism.)

3) But if the cognitive processes in the hunt are purely neural, we seem forced to make the dynamic hunting of the lions a rigid structure like the static kind bees generate by their dance, which is impossible because of the ambiguous (or indefinite) nature of lion behaviour without signification.

What we lack and must identify is what signficatory *link* exists in the world for dynamic communication: what physical signification occurs for the super-organism in the unit of the communicated. For our assumption must be, in the scientific disposition, that while the causal efficacy of the lion hunt lies in purely neural processes, without signification we cannot see how the unit of the communicated is established – how the communicative loop is closed. For both causality and communication are mandatory.

At this point we introduce the biological notion of *brain-sign*. While bees need not signify to others the cause of their *own* behaviour, for causation lies in the static nature of the collective process, lions must signify the cause of *their own* behaviour *as an explanation of their own actions*. It is precisely because of the dynamic nature of the unit of the communicated (i.e. the *one* physical entity which the super-organism is) that an explanation in each organism is required – for its actions may or may not be correct in relation to the

---

<sup>86</sup> For example, Hurley says: 'In goal emulation, you observe another achieving a goal by certain means, find that goal attractive, and try to achieve it yourself. Monkey A may use a tool to achieve an attractive object, leading Monkey B to acquire the goal of obtaining a similar object' (3). What Hurley does not tell us is what 'attractive' means in any physical-functional sense, and she consistently confuses, or does not distinguish observation (or perception) with neural activity supposedly equivalent, presumably on the supposition that the accumulated subpersonal simply becomes the personal, *à la* Dennett, whether isomorphic or not.

<sup>87</sup> Thus the notion of *mirroring*, to which we will come, is confused in the literature. Michael Arbib states, *vis-à-vis* language evolution, 'I will use the term *language readiness* for those properties of the brain that provide the capacity to acquire and use language. Then the hypothesis offered here is that the "language-ready brain" of the first *Homo sapiens* supported basic forms of gestural and vocal communication (protosign and protospeech) but not the rich syntax and compositional semantics and accompanying conceptual structures that underlie modern human languages' (2005; 22). What needs explaining first is how there *is* a brain capacity of supporting a presupposed *conscious* gestural and vocal communication.

other organism, resulting from the potential ambiguity (or indefiniteness) of what is being responded to. Explanatory signification *is* brain-sign. However, no release from determinism occurs in the communal behaviour or brain-sign itself. As such, dynamic communication is entirely at one with static communication – as it must be.

What, in the case of lions, is the content of brain-sign? Brain-sign is the explanation of the cause of an organism's behaviour as it occurs. In the case of each lion, the cause, actual and potential, is the totality of the environment to which it is currently responding.

In the example, the lion's brain is responding to the gazelle, the other lion, and the terrain. Therefore brain-sign will be these features. Since each lion's brain has the structure of the environment before it (in the world) in the hunt process, it can portray that structure as the cause of its immediate causality *in its own brain*: i.e. as an 'image' of the constantly changing terrain, the other lion and the gazelle. That is, the brain's mechanisms produce the image *from* its causal orientation to the world. (The brain does not create the image independently of its causal orientation – by some unspecified means – and then use that as a causal property.) How detailed this image (or pattern) is we could only hypothesise by associating it with the lion's causal capabilities in the environment, and their impact on its behaviour. (This kind of formulation is what pragmatism might have aimed at were it not hidebound by mentalism.)

Thus each lion is welded into the hunt process with the other as one physical structure (genetically precipitated and 'trained' for survival). Its *causal* functioning is complemented, in the unit of the communicated, by a neural *signifying status* explaining its causality. Of course, no lion can look into another lion's brain to see the communicating image. But brain-sign does not exist for the other lion to see, for seeing does not exist in the physical world. Rather it enables the physio-biological process of the unit of the communicated *as* one physical unit, each end (each lion's brain) connected by the electromagnetic radiation between them. In this case, each end (if it is working successfully) signifies the common target, the gazelle. If that is *not* the signification, the hunt will not function successfully because the causal orientation is not unified, as so signified.

Moreover, the image is not some bland appearance, as the notion of consciousness (as knowledge) supposes. (This is addressed in detail later.) Its elements have a time-dependent gradation in causal significance. Chief significance resides in the gazelle as the object of the hunt, with greatest focus on its movements, for therein lies the goal of the biological programme. Varyingly the other lion is significant since its moves must be taken into

account. The terrain too, in its shifting detail, will come and go in emphasis. Thus the elements of the image continually shift in focus as the causal nature of the hunt progresses, from which the image derives. So-called *attention* (i.e. the brain-sign causal focus) derives from the causal propensity of the biological programme in its progress, and not otherwise – i.e. not by some inexplicable mental property.

This mechanism, brain-sign, and its associated physical structure, is a newly identified fundamental scientific neurobiological theory.

The underlying principle is that, for collective behaviour in the dynamic case, causality without communication is a necessary but insufficient condition because explanatory signification is a fundamental component of the structure of that behaviour.<sup>88</sup>

The difference from Hauser is now apparent. He has said that ‘there is no evidence that I am aware of, that communicative signals [of organisms apart from humans] are capable of creating or evoking an image in the minds of listeners, or that signallers have an image in mind when they communicate.’

We agree no images exist in mind. But *mind* has nothing to do with signifying communication. Mind is a dualist presumption (in Hauser’s case arising distinctively with humans) which brain-sign theory discounts.

There is an obvious explanatory continuity with Heidegger’s notion of Being-in-the-world and Being-with with others, for brain-sign is the communicative mechanism associating these notions. But there are fundamental differences too, as we shall come to later.

We encountered the term brain-sign when discussing Peirce (§5.2.1). His use related to language, whereas ours encompasses all the brain phenomena. Moreover, Peirce was ontologically and functionally confused because he retained the role of consciousness (without specifying what it is) in his account of brain-sign. As a matter of historical record, however, I coined the term ‘brain-sign’ before discovering Peirce’s usage, and was delighted to find his precedent with a (possibly) similar intent.

---

<sup>88</sup> We may associate this with feedback in cybernetics in the theories of Norbert Wiener, but we make no further reference to this because cybernetics is founded on different underlying principles.

## 5.4 Advantages of the brain-sign account

Even at this early stage of brain-sign description, we see there are dramatic improvements over consciousness in scientific credibility and explanatory cogency.

### 5.4.1 The explanation of explanation

The first and fundamental improvement is that, unlike consciousness, we have a reason why there should be *any* brain phenomena at all, over and above causal neural functioning. The brain operates as a purely causal organ (not a knowledge organ), but the brain phenomenon, brain-sign, explains the causality of the organism, actual or potential, because only thus can dynamic collective action take place. Since to be causal in the world the brain must respond to the structure of the environment (including the bodily environment), it is in a position, derivatively, to represent that environment. And it represents it, as Husserl proposed, in a 'density' (cf. his inner horizon) because what is represented *derives* directly from the physical response of the organism to the density of the environment in causal terms. (Thus we address the issue concerning explanation in Chapter 1, and revise Churchland's notion of brain theorising, §4.3.)

All mentalist notions can be replaced by this explanatory method, thus establishing humans (and other organisms) as wholly physical in reductive terms.

### 5.4.2 Direct reduction to the physical

The advantage, methodologically, is that signs are physical. There is nothing intrinsically mysterious about brain-sign once its biological rationale is accepted. Although we will not be absolute in specifying the physical constitution of brain-sign (this is both an empirical matter, and relates more deeply to what brain-sign can effect: we will discuss it later), we can say there need be no *correlates* of brain-sign in the brain's substance. For example, brain-sign may be, in part, amassed from the cellular structures of the brain. Analogously, the chameleon changes its appearance by altering its cellular skin.

Both signification and its physical embodiment are biological commonplaces. A crucial advantage of brain-sign *theory* is, however, that it can direct empirical work to a plausible result – by contrast with consciousness theory.

But signs are not knowledge. Thus an explanation of our existence from brain-sign will be quite different from consciousness. Indeed brain-sign, as a signifiatory phenomenon, must be explainable in structure and content by biological classification (as we shall come to) for the direct reduction to physicality.

### **5.4.3 Separation of scientific questions**

By identifying causality and signification as two linked but separate processes, we clarify them as topics in neuroscience. We do not confuse the response of the neural brain to the environment with the opaque notion of knowledge. (We thereby show why the continual change in content of the phenomenon does not incapacitate us in astonishment (cf. §4.1.3) because it is *not* causal knowledge.) On the other hand, we can explore what the brain signifies (or cannot signify) in its explanation of its causality. This displaces the uninformative contrast of consciousness (an unexplainable phenomenon) and unconsciousness.

Moreover, we have a fundamental theory for the brain's functioning, thus founding the possibility of a genuine brain science. For example, what have been termed dream, hallucination, imagination (as images) derive from the same internal causal mechanism as perception. Therefore we have a means for analysing what these states indicate about the organism without resorting to the opaque notion of mind.

### **5.4.4 Evolutionary continuity**

We have chosen lions as the example of dynamic communication because they are non-human. Presumably there are simpler organisms than lions that employ dynamic communication, and evidently there are more complex. Brain-sign theory removes the mystery about which organisms function in this manner, and provides a unifying route through the evolutionary complex.

#### **5.4.5 Dissolution of the inner and outer conundrum**

Both Husserl and Heidegger, by different means, attempted to dissolve the problematic notions of inner and outer (be it empiricist or idealist, §3.1.1), i.e. how does the inner mind get out of the head to know the outer world, or conversely, the world get into the inner mind? Husserl did so with his version of transcendental idealism, Heidegger by the concept of Being-in-the-world. But neither was effective because neither broached the brain phenomenon's physicality. The inner/outer contrast still pervades the literature.<sup>89</sup> Brain-sign theory removes the conundrum because, though within the brain, brain-sign is a function of the brain in the physical universe (contrast, Frith §3.2.1). Thus it is clearer to state that brain-sign is *of* the brain rather than in the brain.

#### **5.5 The integral nature of brain-sign**

In the natural disposition, psychology proposes we possess different faculties: perception, thought, feeling, and so on. Brain-sign theory rejects such classification. The brain phenomenon has one communicative function. But brain-sign has various components as an explanation of the causal propensities of the brain.

In the lion example, we might suppose that in addition to a presentation of the environment to which it is responding, brain-sign may incorporate a neural status which, in the natural disposition, we would term excitement. But no lion knows it is excited, and this component must be regarded as purely neurally indicative. Then again, a lion with a thorn in its paw may exhibit a brain-sign element which, in the natural disposition, would be associated with pain. But no lion knows it is in pain.

The existence of such neural signification as 'excitement' does not derive from the lion seeing the prey, for the lion's brain does not see anything. Thus the theory does not propose that one component of brain-sign can cause another component. That would conflict with its biological function (and be ontologically inexplicable). The sole function of each

---

<sup>89</sup> Merely in titles, for example: John McCrone, *Going Inside* (1999), and Mark Solms and Oliver Turnbull, *The Brain and the Inner World* (2002).

component is to convey to another brain (in principle) a neural indication of the cause of what the lion is doing, or might do.

So we see, by contrast with the proposals of Block and Burge which lack scientific foundation, that there is a way of describing the neural phenomenon without dependency upon language (*pace* Macphail). We stated at §4.1.2 that the brain phenomenon is what we are, not what we can know. It is now apparent (though we have much more work to do) that what we are, ontologically, is a neural sign.

# The Theory of Brain-Sign

## 6.0 Characterising brain-sign

The previous chapter laid the foundations of brain-sign theory. This chapter develops some fundamental detail of the theory by a number of approaches.

- 1) How brain-sign is to be identified as what it is in the physical universe.
- 2) The specification of brain-sign's physical origins and functioning.
- 3) Differentiating the characteristics of brain-sign from consciousness.
- 4) Since brain-sign is derivative of the causal properties of the brain, what its content is and what job it does.

Each of the following sections will indicate the principle approaches addressed. The approaches cannot simply be dealt with in sequence because they interweave. Brain-sign explanation cannot be encompassed in one statement; nevertheless grasping its difference from mentalism – and how the mentalist account is to be replaced – is a constant theme.

### 6.1 The natural and scientific dispositions. [Approach 1]

The advantage of the notion of consciousness is that it provides the medium for knowledge, which means the universe, including ourselves, is available to us as we appear to experience it. In reality the situation must be far more elaborate in explication and grasp.

Accepting the brain's causality lies in its physical states, we are faced, in interpreting that causality, with both complexity and opacity. For if we accept, *via* the theory of brain-sign, that what we are (as brain-sign) *is* the interpretation of the causality of the brain, then *our* interpretive existence does not, as it were, spring from some self-sustaining and self-revealing source (e.g. Descartes' divine source, or Locke's and Husserl's undefined source) lying outside the physical universe (methodologically), which the history of the topic has supposed (and still exists in the notion of the *neural correlates of consciousness*). In our physicalist model, *our* interpretation of neural causality is (and must be) the interpretive mode of the brain of *its own* causality. This is clarified by the function of brain-sign as a

means of inter-organism communication, not knowledge of, in this case, neural causality. We do not *have* an interpretation; we *are* an interpretation – which is not knowledge.

Of course, the notion of consciousness results from (or is at one with) the natural disposition; but the brain phenomenon (as brain-sign) is a mode of inter-organism communication without appreciation of how it happens. (I.e. to us, the communication process is invisible.) The scientific disposition dissolves the simplicity of the natural disposition by grasping the functional biology of the brain's communicative activity. The scientific disposition demonstrates how the natural disposition is reducible to its biological source, not *as* the natural disposition, but as brain-sign. Yet, as said previously, the scientific disposition does not thereby escape the brain's biology, and this must be incorporated in the scientific account. The scientific disposition is still only brain-sign, and thus the limits of science lie within the physical properties of the brain (cf. §5.2.1). There could be no absolute knowledge – even were the idea of knowledge comprehensible. This is compatible with Peirce's pragmatism, though it is not involved.

It is evident, too, that there is no choice in adopting the scientific disposition by contrast with the natural disposition. Given they are interpretations by the brain of its own causal status, both are brain products. We are, in *any* disposition, already determined before we find the conclusions to which 'we' are circumstantially destined – and those conclusions do not aid us in our individual organismic causality.

These points can be summed in a formula. The world is not present to us. We are the presentation of the world to which we are adapted as our (organism's) instantaneous causal orientation. ('We' or 'our' = brain-sign.)

## **6.2 Brain-sign and the virtual array. [Approaches 2, 3]**

Brain-sign is a functional term, founded in a specific scientific framework. It is the state by which the brain signifies its immediate causal relation to the world (that in the world which has caused its causality) in the unit of the communicated. As a functional term it does not specify what in the brain it is, although we can hypothesise. For current purposes we will refer to its physical instantiation as a *virtual array*: 'virtual' because we cannot currently specify it actually; and 'array' because there is a physical status modifiable as the conditions of brain-sign. It is highly unlikely the array is *a* modifiable physical locale or modality, but it

can be referred to as such (for now) because it is precisely the immediate physical condition of its modifiable existence that is important. Setting up these two different terms will allow us to move between functionality and physical instantiation. Since brain-sign is a physical sign there is no conflict between the way we will talk of brain-sign as a function, and its physical instantiation. (The way this works will become clear as we proceed.)

An important aspect of this structure is the elimination of the notion of brain-sign as a *genii* in the lamp that springs into being as something happens in or to the brain, which is how consciousness is imagined. ‘What’, the question is asked, ‘must neurons (or whatever) *do* to become conscious?’<sup>90</sup> From our discussion in Chapter 4, we see this is a question without plausible answer. Brain-sign, on the other hand, is a state of the brain, a condition of neural (and other brain material) physicality. It is generated by the brain in its biological operation, not to reach ‘up’ to a new condition of being. So the chameleon modifies its (cellular) skin under control of the brain as a signifying state (a cellular array). The chameleon has no idea its skin has been modified, and lions have no idea their brains sustain an image of the gazelle – and more significantly in the lion case, that this is the communicative means of enabling the hunt.

A further condition obviates the notion that the brain sustains mental content, particularly memories stored away. We propose the array (whatever it may be) takes on its representational characteristics (as the cause of causality) as an (accumulated) immediacy that disappears as soon as it is no longer applicable. It is not a state, as mental content, that becomes conscious under some kind of impetus, and then subsides back into unconsciousness. Causality lies with the neural brain, and the neural artifice of brain-sign is a transient modality determined by that causality.

The virtual array is not viewed by a subject. Whatever the (so-called) subject turns out to be, to which we will come, it is part of the array, not separate from it. The virtual array fixes the content of brain-sign as an immediate physical singularity. This eliminates any sense in which one consciousness can view another, as in the reflective supposition. For the reflective supposition must be that there are at least two consciousnesses (or perhaps a peculiarly split consciousness), one being what is reflected on, the other reflecting – each

---

<sup>90</sup> E.g. the quantum notion: Stapp (2006), after Roger Penrose. This really relates to the notion of free will (which consciousness is supposed to entail) rather than consciousness itself.

Or the complexity notion: Edelman & Tononi (2000).

Or the endemic notion: Edwards (2005).

Or the frequency oscillation notion: Crick (1994)... Etc.

necessarily sustained by its own subject (to be conscious). As we shall see, by contrast, the *sense* that there is a state of reflection is a functional feature of brain-sign, and the singular virtual array is therefore appropriate as its physical instantiation.

### **6.3 The instantaneous nature of the physical universe. [Approaches 2, 3, 4]**

From the big bang, each instant of the universe is one physical state of being. This physical state is continually changing. Newton's laws of motion (updated by Einstein), and the laws of thermodynamics, for example, insofar as they may be correct, specify aspects of that change.

This does not mean that the universe is determined, as by Laplace's demon: that a hypothetical intelligence who knows the current physical state of the universe could predict all future states. Chaos Theory may mitigate against it (an unresolved topic), and there is the implication in quantum mechanics of an indeterminacy (for the observer) in micro-physical behaviour. But regardless of the ability to predict, what is meant is that, being physical, the state at time  $t_n$  is followed by that at  $t_{n+1}$  inevitably, i.e. by whatever physical conditions do obtain. This is our assumption.

Its significance is the application for brain-sign theory by contrast with consciousness. As brain-sign, we are part of the universe in its instantaneous physical being. *We* (as mental subjects) cannot look forward or back: we cannot consider memories of our past; we cannot imagine our future. *We* (as brain-sign states) are an explanation of the brain's causal relation to the world now. The brain's biological operation lies in the instantaneous nature of its causal relation to the world – for survival and reproduction.

Yet we appear to sustain conditions in which we are *not* locked in an immediate present. But that does not result from our mental ability to roam past and future. It results from the brain because, at such an instant, as well as being causally reactive to the immediate situation in the world, the brain is also active on states of itself which are not of the immediate situation. This does not mean the brain is acting *in* the past or future. It can act only in the present.

Our supposed ability to remember or imagine, even to imagine how our past and therefore present and future could be different, does not remove us from the instantaneous being of the physical universe. Our supposed remembering or imagining, as with all other

states of brain-sign, are completely determined by immediate processes of the brain which we, as brain-sign, neither access nor control.

For example, our saying to ourselves ‘How would things have been if I had not...?’ in which we interrogate our past which seems available to us, requires us, in the scientific disposition, to acknowledge the processes of the brain, rather than supposing, as in the natural disposition, that we have a mental life in which such *mental* interrogation could take place. It requires us to see that the structures of past and future are fabricated by the brain by processes invisible to us as brain-sign. The result of that fabrication, our apparent contentful self-interrogation, exists ‘merely’ for our organism to communicate with other brains on our brain’s immediacy (though not the immediate situation), whether it does so or not. ‘How would things have been if I had not...?’ exists to be communicated. Even if we would be embarrassed to communicate in fact, the arrival in the physical universe of ‘How would things have been if I had not...?’ is *for* communication, and our (brain’s) resistance to communication is an act of communication, not communicated.

That the brain in immediacy is active on states not of the immediate situation is the brain acting normally. But it is inappropriate to propose, as do Milner and Goodale (§3.3.3), that ‘Visual knowledge allows us to plan actions, to picture the consequences of those actions, and to relive (sometimes with pleasure, sometimes with regret) what we have seen and done in the past’, because the function of brain-sign does not include planning actions or picturing their consequences. Similarly Damasio proposes ‘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; 304) (slightly reconstructed) which it does by ‘the world of planning, the world of formulation of scenarios and prediction of outcomes’ (303). But this supposes the future remoulded from the past results from a magic capacity of the brain to overcome the ‘automated responses’ of the non-conscious – with the corollary that conscious states are *not* automated, whatever that may be supposed to entail.

The brain’s causality results from its neural characteristics. ‘Planning actions and picturing their consequences’ as an addition cannot ‘make way’ for (i.e. *escape*) the causal processes of the brain, which itself has no sense of past and future. *Past* and *future* as brain-sign cannot be past and future considered in terms of *planning* actions (and then presumably

implementing them). They can only be signs of what the brain activates concerning its causality, its causality being entirely other in functional terms.

The notion of consciousness removes us from the immediacy of the physical universe. The notion of brain-sign places us in the physical universe. The notion of consciousness proposes, as Damasio innocently states, that there is *another* world of consciousness existing outside the physical universe and not subject to its laws – though doubtless he would claim he does not intend there to be another world.

#### **6.4 Our dual nature: Causality and communication. [Approaches 3, 4]**

Mentalism supposes we are of two kinds: the mental and the physical. The challenge is to demonstrate that the mental is also physical.

Brain-sign theory supposes we are one kind: the physical. But our physicality is of two kinds: the causality of the individual and collective communication, though they are fundamentally linked.

Concerning causality, mentalism is inevitably unclear. George Rey has said (§3.2.1.2):

‘A moment’s reflection reveals we *know next to nothing* about even such basic activities as perception, thinking, reasoning, language, decision making, motor control, not to mention consciousness, creativity, scientific insight, or morally responsible action’ (emphasis added).

The problem is, humans cannot know anything (cf. §4.1.2).

As to communication, mentalism also claims it has a role. We communicate mind-with-mind. But mentalism cannot demonstrate that mind-with-mind can communicate because it cannot define its communicative terms physically. It cannot, for example, tell us what seeing or hearing are.

Note that our use of the word ‘communication’ in brain-sign theory does *not* mean that brains, interlinked in the unit of the communicated, pass each other information (which has meaning and can be understood by minds).

\*

That the physical properties of the brain are the source of causality for the organism's behaviour is not, in principle, controversial. That does not mean there is a general science for how the brain operates in relation to the universe. No such science exists. This problem cannot be resolved here. Our assumption is that in principle the physical properties of the brain *could* be so classified. To move forward on causality and communication we will consider the state of the art.

In the literature it is assumed that there are neuronal structures that perform the role of face recognition.<sup>91</sup> However, though the physical properties of brain regions respond to the characteristics of faces, they do not *recognise* faces because only conscious humans (or consciousness itself) could do that (cf. §4.2, the Bechtel and Abrahamsen discussion). Here is the Abstract of an article by Paller *et al.* (2003) concerning experiments on so-called face recognition.

‘Rapidly identifying known individuals is an essential skill in human society. To elucidate the neural basis of this skill, we monitored brain activity while experimental participants demonstrated their ability to recognize people on the basis of viewing their faces. Each participant first memorized the faces of 20 individuals who were not known to the participants in advance. Each face was presented along with a voice simulating the individual speaking their name and a biographical fact. Following this learning procedure, the associated verbal information could be recalled accurately [by the participants] in response to each face.

‘These learned faces were subsequently viewed together with new faces in a memory task. Subjects made a yes–no recognition decision in response to each face while also covertly retrieving the person-specific information associated with each learned face. Brain activity that accompanied this retrieval of person-specific information was contrasted to that when new faces were processed. Functional magnetic resonance imaging in 10 participants showed that several brain regions were activated during blocks of learned faces, including left hippocampus, left middle temporal gyrus, left insula, and bilateral cerebellum. Recordings of event-related brain potentials in 10 other participants tracked the time course of face processing and showed that learned faces

---

<sup>91</sup> As an artificial neural network methodology see e.g. Uwechue and Pandya (1997)

engaged neural activity responsible for person recognition 300–600 msec after face onset. Collectively, these results suggest that the visual input of a recently learned face can rapidly trigger retrieval of associated person-specific information through reactivation of distributed cortical networks linked via hippocampal connections.’

The title of the article, and the article itself, refer to *neural correlates* of person recognition. The activated brain regions in the experiments are taken as the brain’s causal response to facial stimuli. But identifying brain regions does not elucidate in what way the brain reacts to facial stimuli *qua* face. By talking of neural correlates (of consciousness), Paller *et al.* suppose *the brain* recognises faces, a functional and ontological mistake, brain-sign theory maintains. Though it might seem useful to specify active regions of the brain, this is severely limited as brain science because, in terms of scientific theory, there is no explanation beyond identifying the regions. The terms ‘neural correlates’ and ‘facial or person recognition’ are unmotivated additions (cf. §3.3.4).<sup>92</sup>

Surely the brain must respond to the elements and architecture of the face. Its physical states alter when they are encountered. In so altering, the brain is readying itself for *causal* (not knowledge) response. To respond accurately to *this* face, the brain’s physical states must assume the appropriate conditions. This does not demand a mental *image* of the face (the visual knowing); it is a reactivation of a previous causal *neural* condition.

In the experiments, no behavioural response is required from the participants beyond recognition. But no apparent behavioural response is still, by the brain, a *causal* response. To generalise, that in the universe capable of being responded to by the brain affects the brain’s neural condition. It does not occur to *us* that this is so, i.e. what is happening ‘below’ our so-called experience. We are normally in the natural disposition as which the world simply appears to be there and we in it. I.e. we assume we have knowledge of the world by what we see, not that we (our brains) have already causally responded to the world from which our ‘seeing’ results.

---

<sup>92</sup> In William Bechtel’s debate with John Bickle, he says the following. “The molecular processes in cells to which Bickle proposes to reduce psychological concepts such as memory consolidation are, on the mechanistic reductionists (*sic*) account, processes within operating components of the relevant mechanism. They do not, on their own, explain the phenomenon of memory consolidation. Such an explanation requires the full range of brain areas involved in memory consolidation and the operations each perform’ (2009; 35). But neither actually address what psychological states *are* (as distinct from the behaviour they cause or are associated with), which therefore means their approach is well short of what needs to be explained. Further discussion of Bechtel’s approach occurs in Chapter 8.

To illustrate, consider this from Bennett and Hacker. ‘Seeing an ordinary table or chair does not evoke *any* emotional or attitudinal response whatsoever in normal circumstances’ (2003; 40). Even though they are writing about brain science, Bennett and Hacker do not express the functioning of the brain. They simply absorb the extraordinary fact that ‘a seeing’ has occurred, without questioning either its neural or manifest significance.

The target of their attack is *qualia* or qualitative feel as posited by Block and Searle. The only mental state occurrence, say Bennett and Hacker, is the object seen (table or chair), not an additional ‘emotional or attitudinal response’, as Block and Searle claim. But of course, with both pairs of protagonists the underlying assumption is we do *see* objects.

We remove the oversight of both positions *via* the Paller *et al.* experiments. An object in the world precipitates our ‘seeing’ fair hair as opposed to dark hair. The neural causal condition is not, however, a *seeing* of fair hair. The brain cannot see fair or dark hair. But brain-sign, which the brains of participants generates *from* the causal condition of the brain in relation to the object, now communicates fair as opposed to dark hair. This is the *communicated* status of neural causality with the brains of the experimenters. But it does not occur to the experimenters (or Bennett & Hacker or Block & Searle) for they are in the natural disposition.

Indeed, what is unidentified in the experiments, and in psychological science generally, is how the experimenters’ and participants’ brains (as physical entities) *can* communicate (neurally) to allow the experiments to be conducted at all. That is brain-sign function.

What is in the world *per se* is not accessible to us as brain-sign. (Kant’s position for consciousness, though not for the same reasons.) Neuroscience (as a theoretical discipline) generally takes it that consciousness *is* the seeing of what is in the world – the recognition ascribed *to* the brain by Paller *et al.* Indeed, Bennett and Hacker propose consciousness provides transparent access to the world as if no work was undertaken by the brain at all.

“‘To become conscious of (or that)’ and ‘to be conscious of (or that)’ are factive, or existence implying – that is, one cannot logically become conscious or be conscious of something that is not the case or is not there’ (2003; 255).

But there may be no *object* in the world we see as such. Indeed, perception (automatic and inevitable access to the world *per se*) cannot be a natural state requiring no explanation.

The state must be explained in a plausible biological model, which consciousness fails to do.<sup>93</sup>

\*

Now we might ask the question (as at §3.2.1): *What* is this fair colour of the hair? How has the brain been able to create it? We seem to *see* the fair colour, and even if it is brain-sign, how is there colour when it can neither exist on the object nor be in the brain?

The obvious mistake *now* is the notion we see anything. In supposing we see, we necessarily posit the object seen: then we have to decide whether the object is in the world or in the mind (or brain).

We have dismissed that which sees, the mental subject (§6.1, and to come), and the object seen. In stating the hair is fair as brain-sign we identify a neural condition which is not seeing: nor indeed do we know we see fair hair. That is, the *sense* we know (i.e. what the natural disposition might refer to as a self-revealing apparentness) must be reinterpreted as part of brain-sign itself (as we shall come to).

Thus we solve Rey's conundrum (§3.2.1.2). 'How come no one else can enjoy my reddish experiences by inspecting my brain?' There is no viewing relation between Rey's brain and the viewer. Rey presupposes consciousness with its knowing (in this case viewing) capacity *in* the brain *and* the viewer who could view its *seen* content. No such capacity, or viewer of it, exists. With brain-sign theory, Rey's question dissolves.

## 6.5 Removing the need for causal conscious representation. [Approach 2]

In the previous section we established differences between brain-sign and consciousness in relation to causality and communication. This may be further illustrated diagrammatically. Here is the consciousness model.

Consciousness:            Object    mental seeing    action

---

<sup>93</sup> Francis Crick (1994) wrote a book titled: *The Astonishing Hypothesis* but then qualified it with a subtitle: *The Scientific Search for the Soul*. Crick's science on this topic, manifest in the word 'soul', is rooted in cultural myth. This applies too, to scientific investigators as Charles Darwin, E.O. Wilson and Richard Dawkins because they hypothesise in the natural disposition. What *is* astonishing to the natural disposition is that our existence is controlled by the brain. Our lives are managed by a physical object to which, as brain-sign, we have no access.

A problem is that the unconscious aspects of the process of seeing are not represented (cf. §2.7, Edelman and Tononi discussion). Perhaps they occur before the seeing, and perhaps they (also) occur below the seeing. Owing to the obscurity of the very notion of seeing, there is no location to which they could be (standardly) applied.

The brain-sign model, by comparison, is the following.

Brain-sign:	Object	neural processing, causal establishment	action
		brain self-interpretation	brain-sign

It is obvious, in the brain-sign model, that the brain must interpret *itself* to arrive at the appropriate brain-sign. There is nothing else it could interpret without resorting to magic. Put another way, in responding to the external world *via* transduced input, all the brain can do is alter its physical states. It has no access to the world *per se*. The illusion of the *theory* of consciousness is either that the world is in us, as our mental life, or that we are in the world because we see ourselves inhabiting it. The illusion is a *conceptual* illusion of the natural disposition. The scientific disposition is not subject to this illusion because it describes the situation successfully (i.e. in terms of the physical universe). It takes the *apparent* veracity of the natural disposition as the *function* of the brain phenomenon. The content of the brain phenomenon, i.e. brain-sign, is not an illusion because it fulfils its biological function.

The brain responds to the world for survival and reproduction. There need be no causal representation of the world in the brain. All there needs to be is the capacity for self-modification to act. Brain science has to establish how this happens.

Current brain science, however, consistently insists that causal representation and (e.g.) sensation need to be accounted for in the brain because *we* seem to sustain and act on them. The brain-sign model removes that problem.

How close have we been hitherto to such a possibility? Andy Clark says: ‘The internal representation of worldly events and structures may be less like a passive data structure and more like a recipe for direct action’ (2001b; 88). But why suppose there is an internal *representation* of worldly events and then propose it is a recipe for direct action? If the brain’s function is direct action it hardly needs to represent worldly events and structures. The events and structures are already there in the world for direct action to be a response by

the brain's causal capacities. Clark, in this section of his book, moves closer to the likely scenario.

'Perception is itself often tangled up with possibilities for action and is continuously influenced by cognitive, contextual, and motor factors. It need not yield a rich, detailed, and action-neutral inner model awaiting the services of "central cognition" so as to deduce appropriate actions. In fact, these old distinctions (between perception, cognition, and action) may sometimes obscure, rather than illuminate, the true flow of events. In a certain sense, the brain is revealed not as (primarily) an engine of reason or quiet deliberation, but as an organ of *environmentally situated control*' (95).

We may wonder why the 'sometimes obscure' and 'In a certain sense' are not an acceptance that brain science should reject the notion of consciousness and mind altogether. Clark does not do that. The title of his book is *Mindware* – an indication that his philosophy of cognitive science, despite some diversions, is still gripped by the natural disposition.<sup>94</sup> Indeed the problem remains if we ask: What is the point of an 'inner model' (i.e. consciousness) as an *addition* to direct causality?

One example Clark uses to support his action thesis is *mirror neurons*. Referencing Di Pellegrino *et al.* (1992) and Rizzolatti *et al.* (1996), he describes how these neurons in the ventral premotor cortex of monkeys are active both when the monkey observes an action of another, for example grasping food, and when performing the same action. He quotes Marc Jeannerod as saying that 'the action...to be initiated is stored in terms of an action code, not a perceptual one' (1997; 191). But again, there is confusion in Clark's words because he suggests the monkey *observes* an action of another, which implies (by a tacit mentalism) that there is a monkey 'person' observing (consciously) when no such event in the physical universe takes place. All that occurs is the status change of the neurons themselves. Indeed, it is only *because* no such (observing) event takes place that the plausible *function* of mirror neurons in the physical brain can occur.

(An industry is growing over mirror neurons, and mirroring generally (cf. §5.2.2.7f); but inevitably it is overgrown with mind concepts which obscure (Clark's own word) the physical nature of the neural activity.)

---

<sup>94</sup> However, how the brain operates causally without mental content is a topic under current discussion. See e.g. Peter Machamer, particularly 2009, p. 173. "Suppose the atomic elements (individual axon or dendrite processes, say) don't have content but respond differently to different inputs connected directly or indirectly to environmental influences..." etc.

### 6.5.1 The connection between brain models

We can now directly associate the diagram of the brain-sign model just given with our derived model C (§3.3.3) from the Milner and Goodale two visual streams model of perception. The bottom route from the ventral stream in the latter, which does not influence the input of the dorsal stream, we have specified as going to the ‘brain phenomenon’. The top two routes (‘immediate action’ and ‘deliberative action’) are what is termed in the brain-sign model ‘neural processing, causal establishment’, and lead directly to action. Thus the bottom stream in C is the visual information passing to ‘brain self-interpretation’ in the brain-sign model, and thus to brain-sign. That is, for ‘brain phenomenon’ now read ‘brain-sign’ *via* ‘brain self-interpretation’.

(For our purposes, the debate as to whether Milner and Goodale have identified two separate streams, or aspects of a single stream, is not relevant.)<sup>95</sup>

### 6.6 Introduction to two fundamental elements of brain-sign:

#### **categories-of-the world & categories-of-interaction. [Approaches 3, 4]**

Of course, ‘brain self-interpretation’ is a *façon de parler*. No interpreter in the brain renders brain-sign from causal states. We assume an intrinsic neural mechanism. The word ‘interpretation’ has been used (descriptively) because a transformation of causality to brain-sign has occurred by that mechanism, each having a different functional status. However, the content of brain-sign *is* a guide to the information biology communicates about the cause of the brain’s causality.

Consider again facial recognition (from §6.4). The brain reacts to a face definitively when neural alteration in relation to the elements and architecture of the face assume a sufficient overall condition – a previous condition. The causal relation to the face is then interpreted by the brain to create the brain-sign result, and it can do so because the causal

---

<sup>95</sup> David Milner has said, in discussion, that it was not intended that the ventral stream be influential by being conscious. Unfortunately, from the Milner and Goodale texts, this is the inevitable interpretation. Indeed, if this is an incorrect interpretation, it is not possible to work out how their notion of consciousness is supposed to function. This does not invalidate the significance of their work.

Marc Jeannerod, who disputes their theory, clearly makes the same assumption. ‘The ventral pathway underlies visual perception, i.e. the conscious identification and recognition of objects’ (2006; 7).

relation has established what the world *is* that it is reacting to – not the world actually, of course, but as causally significant to the organism.

In the Paller *et al.* experiments the participants are not in a genuine social situation, which limits the causal orientation of the brain to the faces presented. E.g. participants do not have to interact with the individual whose face they see. I.e. the context of the experiment directly influences the participants *and their brains*.

But what we are moving towards, in the brain-sign model, is the way of avoiding the confusion about what the brain is doing. In a concluding section of the paper by Paller *et al.* titled ‘Person Recognition and Memory Retrieval’, they state:

‘Person recognition is a prime example of the type of remembering that patients with memory disorders find difficult. It depends on the storage of diverse sorts of information in multiple cortical regions. Neuropsychological studies have begun to delineate the distributed network of brain regions responsible for storing and retrieving such memories. Neural correlates of person recognition in neurologically healthy individuals, as identified in the present investigation, can corroborate these findings and help to specify the distinct roles of each region. Further research in this area can thus enhance our understanding of how complex interactions among brain regions can swiftly mediate memory retrieval in response to an appropriate stimulus cue such as the face of a known individual’ (2003).<sup>96</sup>

However, the ‘stored’ states of the brain in relation to a face concern the brain’s causal reaction to the face, not its features as so-called memories (cf. §§2.9 & 6.2). Fair hair is not stored in the brain as some kind of unconscious (whatever that may be) representation to be recalled in the appropriate conditions. The brain’s *causal* orientation to elements of the universe are located in ‘multiple cortical regions’ and aggregated into a complete causal effect. The appearance of a ‘face’ (brain-sign) is a consequent interpretation facilitating communication concerning that causality.

How could we corroborate this? One way is to appreciate that brain-sign entails (what the natural disposition would term, somewhat coarse-grained) a judgement on its content. Faces are not passively encountered. Fair hair is judged in itself, and so is the rest of the face. Faces are: attractive, repulsive, confidence inspiring, dubious, and so on, to the subtle point

---

<sup>96</sup> No page number is given because the online version is referenced: para. 2 in section.

of linguistic unspecifiability. These ‘judgements’ are not derived from a viewed representation of the face (a so-called percept), but arise directly in relation to the face itself. They explain how the causal brain *responds* to the face. They are biologically grounded, and reducible ultimately to survival or reproduction dispositions in the fabric of the brain.

Two people looking at a face can communicate (in the dynamic structure) because their brains sustain an adequately equivalent brain-sign state of the face. But each, if asked, will qualify the face with a different judgment, because each brain responds differently. Both image and judgement, as brain-sign, are the means of brain communication – but, as such, are unbeknownst to the individuals because they (i.e. their brains) are comprehensively rendering ‘them’ as the natural disposition.

Facial judgement relates to factors of personal interaction. But we can associate the notion of judgement, in a case Heidegger might have chosen, with a response to a fork. The brain encountering a fork (from ‘learned’ acquaintance) does not *recognise* it, but relates to it as an object to be used in a certain context. Forks are not stored in memory in the brain (*à la* Paller *et al.*), but the overall causal relation to a fork, as brain-sign, will emerge as the image of a fork *along with* the *sense* of its use, even if at this moment it is not going to be used. Its *sensed* use is the biologically grounded causal orientation of the organism. In other words, the response to a face and the sense of use of a fork have a common bio-functional character because they both represent the organism’s causal orientation to the object *per se*.

\*

From this discussion we may conclude that it is unnecessary to suppose, as conscious cognition must, that inside the head are two interpreters. The first interprets input from the world to create the conscious image,<sup>97</sup> the second interprets the conscious image to decide what to do about it.

If the brain reacts to the world directly, and as a neural network (although this cannot be the only physical means of causal response), there is no need for either kind of interpreter,

---

<sup>97</sup> As with David Marr’s (1983) computational: primal sketch, 2½D - 3D images.

nor the associated mental processes of cognition, *as* e.g. understanding or knowledge, from which action follows (as in diagram, §6.5).<sup>98</sup>

Let us, therefore, make a further clarification. We are not yet going to discuss language. However, the notion of judgement (about e.g. faces) to which we have referred has been associated, historically, with language. On the other hand, fork usage, which seems to be of a similar neural-responsive kind, is obviously not a judgement but a *sense* of use (cf. 6.4).

Our real target (what we are leading up to) concerning the so-called judgement about faces is not verbal but rather how we (would) respond *behaviourally* to the face, which subsequently may be expressed verbally in terms of ‘attractive, repulsive, confidence inspiring, dubious, and so on’. Prior to any such verbal judgement, there is a behavioural one (at least as *sensed*), and that is what is intended. Similarly for the fork. We sense, in viewing the fork, how we would behave with it.

The face/fork behavioural *senses* are not added to the image: i.e. we do not perceive what is in the world to which a behavioural sense is then formed. They co-occur because both are necessary aspects of the brain’s effective response to the world as communicated (*pace* Bennett & Hacker’s seen table and chair, §6.4).

The image of the face, for example, is already different between two observing people, and *our* behavioural sense is the specific orientation of *our* causal neural status. This is not fixed, but may change along with the image. So it is we can ‘see’ things at different times in different ways – because our causal relation to the object has changed. (There is a direct comparison here with Husserl’s transcendental idealism (§5.2.2), although brain-sign is neither transcendent nor ideal.) On the other hand, while two people will ‘observe’ a fork from different angles, the sense of its use (i.e. the behavioural response) still occurs in both. Thus neither change nor difference prevents the image and usage *sense* functioning in the unit of the communicated.

But they can also be the modality of communicative disagreement. The natural disposition supposes we disagree upon what we see *as* the world. This is incorrect, and again, *vis-à-vis* Husserl *vs.* Descartes about mentality (§5.2.2), it confuses world and (not mentality

---

<sup>98</sup> We use the expressions ‘network’ or ‘neural network’ to summate neural conditions generally (i.e. at all ‘levels’ of functionality). In the literature, the expression pathways or neural pathways are sometimes used. Occasionally we will use the expression neural network (+) to emphasise the generality.

but) brain-sign. For the disagreement derives from the difference in causal orientation, which is what is being communicated.

\*

The face or fork themselves, as brain-sign, are examples of what we term categories of the brain-sign world, or simply *categories-of-the-world*.

The general term we introduce for brain-sign content as behavioural or usage sense is categories of organismic interaction, or simply *categories-of-interaction*. What is pointed to is the *how* of the organism in its interaction with the world. We will discuss these classifications more extensively later. But two preliminary points concerning neural reduction should be added here.

(1) In the hunting lions example in the previous chapter, we referred to the gazelle image. The gazelle image is a category-of-the-world. But the word 'image' is really incorrect because it would imply something viewable, which categories-of-the-world are not. Categories-of-the-world are neural states; their ontology is not to be confused (as with Locke (§3.1.3)) with the supposedly see-able world.

(2) Agreeableness or attractiveness (before verbalisation), felt in the natural disposition concerning a face, has been taken as a mental state and irreducible to the physical. But on the contrary (in the scientific disposition), what has been taken as (e.g.) agreeableness is actually a category-of-interaction as a neural sign denoting *directly* a behavioural response (at least potentially).

Thus signification and what is being signified are both entirely physical. In Peirce's phrase: there is absolutely nothing more in it.

## **6.7 The content of brain-sign. [Approaches 2, 4]**

Gerald Edelman has coined the felicitous phrase for perception: *The remembered present*. Although Edelman & Tononi's conception of consciousness is unclear and involves suppositions he does not explore (§2.7), he has pointed to a characteristic which requires investigation.

In one place he states:

‘The ability of an animal to connect events and signals in the world, whether they are causally related or merely contemporaneous, and, then, though reentry with its value-category memory system, to construct a scene that is related to its own learned history is the basis for the emergence of primary consciousness.... The ability to construct a conscious scene is the ability to construct, within fractions of a second, a remembered present’ (2000; 109).

What Edelman & Tononi point to is that everything that is (not consciousness but) brain-sign is, to a great extent, already established.<sup>99</sup> In ‘learning’ to respond to the world, the brain self-programmes its causal conditions. These causal conditions, in their assembled constituents, endlessly reoccur until the organism ceases to exist. They are the foundation of brain self-interpretation as brain-sign.

The fair hair the participants ‘see’ in the Paller *et al.* experiments is not an object seen now. It is not even a novel or immediate interpretation of the actual conditions in the experiment. Mostly, every aspect of the brain-sign category-of-the-world is a reoccurrence from previous events triggered by current input. This is what Edelman & Tononi intend by the word ‘reentry’.

The idea this is not so, that we do see what is there, is founded, as the notion of consciousness itself, on the supposition that we can appreciate the intrinsic nature of our own (immediate) state because it is available for us to interpret as such. Because Edelman retains the notion of consciousness in his account of the brain phenomenon described as re-entered, consciousness (i.e. *knowledge* of the immediate world) is devalued to the point of superfluity. While Edelman & Tononi’s diagnosis of the construction of the brain phenomenon is compatible with the approach here presented, the notion it is consciousness (with its value-category memory system, which cannot by definition be physical) is rejected.

In the natural disposition we do not know that almost everything we see has already been constructed years ago (much in infancy), because we had no sense of the construction as it took place. While as programmed we take this endless reoccurrence as *our* immediacy in the world, its biological function is for immediacy *itself* to be communicated – as brain-sign (cf. 4.1.3.).

---

<sup>99</sup> Although the *remembered present* is a felicitous phrase, the concept is not new with Edelman.

Our *causal* immediacy lies in our neural relation to the world which, although largely pre-constructed, necessarily must account for what is particular or even original *now*. But since brain-sign, as categories-of-the-world, hardly sustains any content related to our causality (it is the neural sign of the brain's cause of causality), it is to an even greater degree 'mere' reoccurrence. The Paller *et al.* experiments participants may never have seen this particular face before, but its assembly is almost totally derived from previous faces, and more primitive elements including hair shape and colour. Indeed, if there was anything radically different about the face, for example barbed wire from the top of the head, there would be a startled response, resulting from the potential causal reaction of the participant's brain. But even the barbed wire would be pre-constructed, though it might take a little longer to 'appear', because it 'should not be there'.

### **6.8 Control and duration. [Approaches 2, 3, 4]**

The brain is a control system (§6.5). Its function is to operate the organism in the world (including internal environmental control). As organism, we are a continual series of bodily action programmes running from initiation through the action itself to completion, if successful.

One might suppose that any text about the brain would start here. But this is not so because of the supposed requirement to account for consciousness and its (e.g. knowledge) characteristics. For consciousness is set over *against* the biological reductive domain, in explanatory terms. Indeed, the reversion from behaviourism to mentalism that Chomsky initiated, to which we have referred, was no move forward. Behaviourism's model (stimulus-response-reinforcement) was too simplistic; it lacked a means of mental reduction to the physical/biological, and it had no concept of the function of the brain we have termed brain-sign. But with Chomsky's reinstatement of mentalism (giving back a legitimacy to consciousness), the science of the human being returned to myth and obscurity.

\*

Although the brain is a causal organ operating under describable laws, even if as yet unknown kinds of laws (for brain science lacks these fundamentals),<sup>100</sup> and as we have said,

---

<sup>100</sup> This is not to say that there are not theories, as large systems dynamical theories. But these do not as yet constitute an agreed *science* of the brain. John Bickle would not agree. 'As long as philosophers (and cognitive

state  $t_n$  is followed by state  $t_{n+1}$  inevitably, brain states *endure* because the actions of organisms themselves endure. Lifting an arm or arranging to go to the seaside require the neural brain to sustain protracted states, even with prolonged lapses of relevant activity before they are resumed.

Thus the priority of action over (supposed) mentality renders Kant's kind of analysis redundant. He states: 'When I seek to draw a line in thought...obviously the various manifold representations that are involved must be apprehended...one after the other. But if I were always to drop out of thought the preceding representations (the first part of a line...), and did not reproduce them while advancing to those that follow, a complete representation would never be obtained' (1933; 133).

But the action of the brain in the drawing of a line does not *depend* upon (mental) representations that must be sustained (as is the case in the consciousness model, §6.5). Brain-sign, as categories-of-the-world (e.g. the line), results *from* the action and will be as sustained or developed as is relevant to the action.

Yet how, one might ask, can the brain want to, plan to, and execute going to the seaside without these being the mental states of a knowing human being? How could brains move us about without there being some kind of intention? Surely it is because we *want* to go to the seaside, that we will *feel* relaxed and invigorated, that we will *enjoy* the spectacle and atmosphere, that we will be able to *see* the sea, and perhaps visit some interesting sites and *learn* about the history, that we do go.

(We have Nietzsche's commentary here. The vanity of spirit is to suppose it is 'an end in itself' (1961; 61-62) – that *we* want, feel, enjoy, see, learn. And again, Sartre's notion that we overcome (falsely) the evanescence of consciousness (or it overcomes itself) by taking it there is a substantial state as seeing the sea, in that *we* see it, and *it* (what we see) *per se* exists *as* substantial, enduring *and* causal (1957; 566).<sup>101</sup>)

---

scientists) ignore developments in cellular and molecular neuroscience, they will continue to either voice anti-reductionist worries or look for mind-brain linkages in the higher-level reaches of cognitive science. But molecular neuroscientists have developed experimental practices that bridge the behavioral to the molecular pathways *directly*' (2006; 414). Unfortunately the issue is not the linkage between behaviour and brain, but as Bickle himself (disapprovingly) says 'the higher linkages of cognitive science'. On the other hand, Bickle does not explain inter-neural communication which, as brain-sign, removes the reductionist problem altogether.

<sup>101</sup> Quoting Ronald Hayman, 'This reminds us of Sartre's 1939 expostulation to Pieter [a wartime colleague in the military]: "I know in my bones that people aren't: they do." At this level Sartre's thinking is anti-Heideggerian and his emphasis on action accords with his insistence on the insubstantiality of consciousness,

On the contrary, human (or neural) self-interpretation as consciousness (the natural disposition) has led to three fundamental mistakes. Firstly, as brain-sign, we are not an immediate acquaintance of ourselves with what *is*, but (largely) a neural reoccurrence which exists to communicate the immediacy of our organism to others. This discounts the notion that we as consciousness are ‘an end in itself’.

Second, the seeming enduring of this sea before me ensures for me that the sea must exist as the same sea as there was before *as I saw it*. This is the rationale within Bergson’s notion of duration, distinguished from the scientific concept of time, and derivative of the powers of consciousness. He says:

‘In our duration – the duration which our consciousness perceives – a given interval can contain a limited number of phenomena of which we are aware. Do we conceive that this content can increase; and when we speak of an infinitely divisible time, is it our own duration we are thinking of?’ (1991; 206). No. ‘We seize, in the act of perception, something which outruns perception itself, although the material universe is not essentially different or distinct from the representation we have of it’ (208).

Brain-sign theory states that duration results, not as or from consciousness, but *via* a sustained neural engagement with the world which, (neurally) interpreted as brain-sign, must also be sustained. This model is not to be differentiated from the scientific concept of time. Moreover, the duration of brain-sign says nothing about the credibility of its content *vis-à-vis* the world.

For the third mistake is that there is an equivalence between what we see and what is out there (the Lockean fallacy). Since brain-sign is a neural interpretation of the cause of the brain’s causal status, the cause being a ‘mere’ modification of the brain’s physicality, what we ‘see’ must be a neural *invention*. So it is we ‘see’ colour (apparently). But this condition of invention must apply, too, to *all* brain-sign elements. We (as consciousness) are not an alternative but essentially similar world to the actual world.

What causes us to go to the seaside and enjoy ourselves has nothing to do with planning to go or enjoyment. The brain creates wants, plans, enjoys, etc., as the natural

---

which he defines in *L’Etre et le Neant* as “the being which has in its being consciousness of the nothingness of its being” (1986; 190).

disposition has them, as the means of neural communication. While we suppose our 'mental states' are 'ends in themselves', we lack a scientific version of the human being.

But it is evident, from this analysis, why and how we live stories (cf. §6.10).

## 6.9 Subject, self, ego. [Approaches 3, 4]

Historically, the subject is an unstable and problematic category. There is the Lockean forensic approach; the Humean (self-confessedly dubious) eliminativist approach; and the Kantian foundational transcendental approach. Kant's model seems well worked out for, after all, it is *I* who am aware of being conscious. However in the modern literature, the subject and their consciousness are without a generally agreed underlying theory. This is made even more obscure once the physical brain is introduced.

\*

The foundational significance of the subject of consciousness, with Kant, is expressed by Heidegger. As already quoted (§2.2f): '[Kant] continually inculcated that the ego is not a representation, that it is not a represented object, not a being in the sense of an object, but rather the ground of the possibility of all representing, all perceiving, hence all the perceivedness of beings and thus the ground of all being.' So foundational is it, indeed, that it cannot be 'accessed' by us.<sup>102</sup>

In the A edition of the first *Critique* (1933: 343) Kant says:

'We are unable from our own consciousness to determine whether, as souls, we are permanent or not.... For since the only permanent appearance which we encounter in the soul is the representation 'I' that accompanies and connects them all, we are unable to prove that this 'I', a mere thought, may not be in the same state of flux as the other thoughts which, by means of it, are linked up with one another.'

In the B edition this becomes (380):

'I think myself on behalf of a possible experience, at the same time abstracting from all experience; and I conclude therefrom that I can be conscious of my existence even

---

<sup>102</sup> Cf. Searle's reprise, §2.5f

apart from experience and its empirical conditions. In so doing I am confusing the possible *abstraction* from my empirically determined existence with the supposed consciousness of a possible *separate* existence of my thinking self, and thus I come to believe that I have *knowledge* that what is substantial in me is the transcendental subject.'

Although there is no agreed theory in the literature, the subject is tacitly foundational for consciousness (e.g. Nagel, 1974; Damasio, 1999), and enduring. This position will be termed *egocentrism*.

Brain-sign theory, by contrast, states that the role of the brain phenomenon is bodily not mental. The subject (self or ego), as an aspect of brain-sign, must be part of the virtual array. Its ontological status, therefore, is not as a privileged entity. This position is termed *neuro-centrism*, set in the context of a biological model, which is termed *bio-centrism*.

Since each instance of brain-sign is unique, 'the subject' cannot be continuous or enduring. We might sense, in the natural disposition, that we, the subject, do endure. But this cannot be established in practice (as Kant says), and according to the theory, we (the 'I') cannot know we endure, for we know nothing. We certainly could not know 'ourselves' as the 'I', for in the virtual array, *being* the instantaneous subject is precisely the same, ontologically, as being this tree I 'see', i.e. it *is* a representation (contra Kant) which is then gone as my body turns around. The difference is that what the subject *refers to* has not gone, which is why it still occurs.

In brain-sign theory the 'I' refers to the body of the organism. (It is not the self-presence to itself of the mental 'I'. It does not tell us we exist, which would be another mentalist tautology.) This is in sympathy with, though not identical to Schopenhauer ('The world is my will'), for we replace the anthropomorphised term 'will' with 'casual orientation'. When I 'see' a tree, brain-sign content has the tree present (as category-of-the-world) but also the 'I' as (bodily) response to the tree. The cause of the causality of the organism is the tree and the organism itself (the 'I') *as my 'seeing' of it*. This is how the 'I', as representation, is reducible to the brain. (As Nietzsche said (after Schopenhauer): The 'body and its great intelligence...does not say "I" but performs "I"' (1961; 61-62).)<sup>103</sup>

---

<sup>103</sup> (1969; 4) Schopenhauer, however, did not propose the 'I' represents the body itself. 'The subject does not lie in space and time, for it is whole and undivided in every representing being' (1969; 5).

By contrast, we point again to the failure of Bennett and Hacker to grasp the function of so-called seeing (§6.4). ‘Seeing an ordinary table or chair does not evoke *any* emotional or attitudinal response whatsoever in normal circumstances.’ This takes no account of the brain’s role. ‘Seeing’ is the explanation (in the unit of the communicated) of the causal response of the organism (‘I’) to the tree (category-of-the-world), and in any further causality, e.g. the sense of desire to go and touch it (as category-of-interaction), the ‘tree’ is contextualised by that sense, and possibly the following bodily movement. That the ‘I’ occurs more often than the tree does not make the ‘I’ more permanent. It is just as evanescent for it is always contextualised to that to which it (the organism) is responding. We *as* the ‘I’ never notice any difference *in* the ‘I’, for *we* notice nothing by not, *per se*, existing.

Indeed, the *sense* of ‘I’ or subject in brain-sign is a category-of-the-world as is the tree. There is nothing strange about the ‘I’ representing the body: after all, the tree we ‘see’ is no more like a tree, ontologically, than the ‘I’ is to the body. But the body the ‘I’ represents is not the body that appears in brain-sign: it is the casual brain-body-organism we do *not* ‘see’ as brain-sign. Which is why it is not represented in the terms of the tree, i.e. as an ‘image’.<sup>104</sup>

The description of the brain-*sign* structure could be seen as similar to the syntactical structure of a proto-language (by contrast with consciousness). The ‘I’ as subject (which is the body); the world as object; and the linking relation as the verb.

---

<sup>104</sup> This may be contrasted with Thomas Metzinger’s approach. Since, in principle, Metzinger intends to be biological and brain-orientated, there can appear to be sometimes a similarity of statement. E.g.:

‘At least for all conscious beings so far known to us it is true they neither *have* nor *are* a self. Biological organisms exist, but an organism is not a self. Some organisms possess conscious self-models, but such self-models certainly are not selves – they are only [!] complex brain states. However, if an organism operates under a transparent self-model, then it possesses a *phenomenal* self. The phenomenal property of selfhood as such is a representational construct: an internal and dynamic representation of the organism as a whole to which the transparency constraint applies [i.e. it does not know what it is]. It truly is a *phenomenal* property in terms of being an appearance only’ (2005; 3). But at the same time the phenomenal self-model ‘is not a thing. It is an integrated process’.

However, Metzinger depends on the notion of supervenience for consciousness. ‘Consciousness not only locally supervenes...it locally supervenes on parts of single organisms’ (17). But supervenience can be a means of talking as we wish without identifying the ontology (or reality) of what we are talking about. In Metzinger’s case consciousness itself, which is not addressed. Without clarifying what consciousness is, therefore, we ask what is meant by ‘self-models are not selves – they are only complex brain states’? Are the elements of self-models thereby causal? How? What is meant by ‘a *phenomenal* property [is] an appearance only’? An appearance to what, to whom? And what *is* phenomenality? Is it causal? And what kind of non-thing is a phenomenal self-model which is a process? Unfortunately Metzinger’s position is no clearer than mentalism itself; yet (presumably) he intended it as science.

Also Metzinger (2003).

Although we will not address it yet, this structure could be in place with other creatures because the ‘I’ is not a mental subject. But clearly in man (and perhaps some other creatures) it is at a different level. (Cf. §§5.2.2.1 & 5.2.2.7)

### **6.10 Memory and imagination reinterpreted as physical states. [Approaches 2, 3, 4]**

What of our past and future? Can we consider our past and imagine our future? (We have touched on this at §6.3.)

We have characterised brain-sign content as a neural fabrication, derived from the causal status of the brain. The brain can only operate in the present. Therefore the notion there is a past and future (latent and to be evoked) in the brain is a misconception: *we* have no past or future. For we, as either brain-sign as a whole or the ‘I’ itself, possess no ontological status ascribable *to* the past or future. On the other hand, as said, almost everything we are as brain-sign is a prior construct, so we need to get clear about what is involved.

Evidently the brain constantly works through its own states in two ways. Without knowing what it is doing (as we suppose we do in the natural disposition, cf. Dennett §4.2), its processes involve *repetition* and *reformulation*. They may be operationally linked. They are an aspect of the brain’s functional/adaptational self-organisation, and can occur without instigation by the immediate situation. But the brain is not bringing up memories or futuristic possibilities *for us* as mental subjects.

The brain organises or reorganises its *causal* conditions – its neural network structures, and other types of material. What the repetition and reformulation are *for*, in their functional/adaptational operation, and whether they work successfully, is a topic for brain science. We will not address it in detail. The follow on, with which we are concerned, is the status of brain-sign *from* the neural repetition and reformulation.

In the natural disposition we ‘take’ memory (i.e. the result of repetition) as of the past. But the contrast with the present is not a differentiation by the brain of what is past actually as opposed to what is present. The brain knows neither past nor present. It does, however, differentiate its causal immediate activity from its causal activity which is not immediate and not reformulated, which signifies as brain-sign of the past. ‘The past’ is not ‘fainter’ because it is a recurrence of an impression which was originally ‘forceful and lively’, as Hume would

have it (1962; 45). It is generally less compelling because the (neural) causality from which it derives is not *effected* by the brain now. This is what the past *is* reductively: ‘the past’s’ neural ontology. The less the brain is operating on the causally immediate, the more ‘the past’ can be current brain-sign.

The brain contains no memories: our parents, our friends, the places in which we have lived do not exist in some latent (i.e. unconscious) form in the brain which we can experience when conscious. It is evident, therefore, that the process of the occurrence of e.g. (supposed) visual memories is replicative of, though functionally different from (supposed) perceptions.<sup>105</sup>

We see why Milner and Goodale’s claim that ‘visual knowledge allows us to...relive (sometimes with pleasure, sometimes with regret) what we have seen and done in the past’ misinterprets neural function (as does Damasio, §6.3). As brain-sign we do not ‘relive...the past’. The ‘reliving as memory’ (in the natural disposition) is a communicative modus of the brain *for others*. The significance of our *neural* ‘past’ is its possible causal self-activation actually. *As is* (i.e. neurally), it is not available as brain-sign. Most of what we ‘remember’ in this self-activation without action, therefore, since it is not communicated (because e.g. no one is there to communicate it to), is functionally superfluous. But it is not epiphenomenal. It must occur now because brain-sign is neural communication from causal activation. The brain cannot switch off its mechanisms because there is no one to whom to communicate brain-sign content.

Memories, as the natural disposition has them, serve no purpose. What role could they have? What is functionally significant is how the brain *causes* the organism to act in a current or future situation which is a (near enough) repetition of a past situation, neurally specified. But these cases are of course *not* repetition in the world (for the world is inevitably different), but functional reoccurrence by the brain in the world (§6.7).

In repeating causal structures now which are not acted upon there is an activation of the immediate ‘I’ in relation to the repeated structure. This is why, as brain-sign, they have the *sense* of being memories of the past (for the ‘I’ of the ‘past’ exists in the repeated brain-sign). Repeated structures activate elements of the then categories-of-interaction associated

---

<sup>105</sup> Although Edelman & Tononi are attached to the notion of consciousness (as we have seen), their view in this respect is similar. E.g. ‘A memory is not a representation; it is a reflection of how the brain has changed its dynamics in a way that allows the repetition of a performance.’ (2000; 95).

with categories-of-the-world, and so invoke the *sense* of ‘past’ ‘feelings’ (etc.) as memories. But no subject remembers *its* past: the entire construct is neurally explainable with the elimination of mentalist concepts.

\*

Let us take this further. Let us again consider, as at §6.3, what is involved in supposedly imaginatively interrogating our past without mentalism. We introduced the phrase ‘How would things have been if I had not...?’ which seems to require, for such a question to be asked and answered, that we *can* interrogate our mental life. It seems that unless we had a memory of the past, and unless we could imagine our present (or future) differently, not only could we not ask and answer such a question, what would be the point of imagination at all – as a mental property?

The neural account replaces the mentalist model of imagining alternatives to ‘if I had not’. Rather than talking of juxtaposed mental states (past, present and future), the brain, in its self-reorganisation, constructs an alternative neural network to that which sustains the conditions that obtain today. Thus the alternative network is that from which the brain-sign ‘if I had not’ is derived.

(1) In what way does the neural network of today’s situation obtain?

The network of today’s situation is the causal condition of the organism’s brain signified (to the degree it is) by brain-sign. Let us say it refers to a financial investment we made. Associated with first making that investment are numerous causal processes resulting in the investment being ours. (As brain-sign, these causal processes may be communicatively represented as, e.g., my decision to invest, the process of making the investment, the watching of the investment going down.)

But why had I made the investment at all? A principal reason (as the natural disposition would have it) is that I thought it would do well. But what would be the point of the investment doing well? It would have improved my financial situation. What would be the point of improving my financial situation? In general terms, I would have expanded my control over my life. Or more specifically, and to make the final reduction, my brain would have achieved more control over the world for my organism, which would have improved its chances for survival – and thus ultimately fulfil its biological *modus operandi* (cf. §6.8). That control must be identified in purely neural terms – a condition for brain science to establish.

Now a mentalist might question this reductive account. They might say: ‘But how could you know that investing positively improves your chances of survival unless you understand financial affairs, and indeed the chances of survival? You can’t do all this by unconscious instinct.’

The answer, however, is quite straightforward. From a neural stance, ‘all’ that is needed is the causal operation of the brain orientated to biologically directed processes, and neural networks that ‘learn’, or are trained over time (the time of a human life from infancy, cf. §6.5). The mentalist questioner (in the natural disposition) confuses what they take to be mental states with brain-sign states. Mental states offer no functional addition to biologically directed and trainable neural network states. But without brain-sign states as part of the causal process in the unit of the communicated, organisms could not have, *collectively*, financial markets in which to invest.

(2) Why does the brain produce alternatives which cause us to regret (‘if I had not’)?

Producing alternatives, which the natural disposition takes as imaginings, do not occur, in Damasio’s words, because ‘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.’

To respond to novel situations generally, the neural brain *de facto* is capable of creating new neural network structures which do not transcend (ontologically) the condition of ‘automated responses’ (a surprising phrase for a neurobiologist to use). How and in what way this is achieved is a topic for brain science. But it does not require conscious imagination.

That a new neural network (causal, but not actively so) is brought into being as a contrast to the situation as obtains today, must result from a reformulation capability that in principle, if not always (or even often), facilitates improved behaviour. (Again, this belongs to brain science in terms of the ability to have control over the world.) We do not learn from our mistakes by undergoing a state of regret (*pace* Milner and Goodale). What the natural disposition takes as regret is a neural invention (category-of-interaction) signifying the *difference* the alternative, as a more survival-fulfilling behaviour, would have accomplished. I.e. we would not have suffered the negative consequence of our actual action. (The

comparison process may therefore be seen as a means of neural self-regulation.) The *sense* of regret occurs whether the alternative would have been more effective or not, and that, of course, is unestablishable – for the organism is not (at this moment) acting in the world actually from which effectiveness could be (neurally) determined. ‘Regret’ exists as a communicative mode for the unit of the communicated as a category-of-interaction, as do the elements of the alternative course of action as categories-of-the-world.

Here is an example. The alternative network might signify as: ‘I should have left it in the bank.’ But clearly if no risk is taken, there is no improvement possible for survival. Therefore regret that the risk was taken in the investment is not signification of a generally more survival-based effective behaviour; it is simply a comparison of behaviour which demerits what *was* done because the result is ineffective. (This does not mean, of course, that behavioural networks cannot operate on some form of statistical likelihood. But that is not our current consideration.)

I do not *know* I suffer regret because I chose the wrong strategy and another would have been better (whatever I say). As brain-sign I *am* the signification of the difference between the two (causal) networks communicated by the sense of regret, and its (possible) articulation. Our *physical* existence as ‘what we did’ is the insurmountable given, and we could have done no other, because the neural networks that caused us to do what we did caused us to do what we did. (This way of expressing the situation demarks the difference between a physicalist account and that of Milner & Goodale and Damasio.)

As causal potentialities, alternative networks now (from reformulation) exist for the future. These may cause us not to ‘make the same mistake’, and we may, in the natural disposition, relate that to our previous regret as we tell it to others. But the determining mechanisms for our actions are invisible to us (as brain-sign). Communicating why we did not make the same mistake may evoke the sense of regret as the cause of the actions we next take: but the sense of regret in fact had no influence. And indeed, our alternative may *not* turn out to be better, and regret will return. For a reformulated network and the sense of regret do not constitute knowledge, though in the natural disposition we may attribute knowledge to ourselves if we are successful (the supposed knowledge being brain-sign as neural communication).

## 6.11 The scientific explanatory power of brain-sign theory vs. mentalism.

[Approaches 2, 3, 4]

Samuel Guttenplan states generally of philosophy of mind: ‘The idea that the mental realm sits on the bedrock of the physical world as described by science, and that we shall gradually, and in our own good time, come to understand the contours of this fit, is by and large the common view of the scientific community’ (1994; 78).

Brain-sign theory offers an entirely different account: it denies mental states.

How, then, do we reduce brain-sign states to their physical actuality, or alternatively, so derive them? The answer is that causality and brain-sign states are expressible in purely biological terms and, *as such* are physically reducible. So far we have specified the brain’s essential condition as causality (not knowledge), and replaced the mental by transient brain-sign states derived from the brain’s instantaneous causal states.

Let us go further. We have given an example of biological reconstruction in the previous section. Regret is not a mental state but the sense of the comparison of one course of action with another (as category-of-interaction), the actions themselves being generated by neural network (+) states. Let us lay this out a little differently.

A series of active network states A caused us to go through the process of an action. But the network states can be activated without actually causing the action (sometimes referred to as being off-line). Associated with the activated (but not active) networks is the neural interpretation which is the sequence of brain-sign states. These constitute what the natural disposition refers to as memories. Since the network states were designed to achieve an effect in their originally active mode, and since other network states B are capable of reacting to the world as the organism’s causal response to A’s result – in this case as a failure *now* in the world to be what the original networks set out to effect – the resulting brain-sign states of A *vs.* B, a *difference* of the two, is interpreted (in the natural disposition) as disappointment. But disappointment is turned into regret (in the natural disposition) by the *difference* of A from the network states C of an alternative action (not activated) with a hypothetical result expressed as brain-sign, which the natural disposition terms imagination. But the difference is not the brain making a *comparison* between states, but rather a neural *oscillation in time* between states A & B and A & C resulting in the neural expression of ‘disappointment’ and ‘regret’ as categories-of-interaction. I.e. we do not hold two states

before us simultaneously as a comparison (as a person might be supposed to do *via* consciousness), but find ourselves moving between them (hence the oscillation) as the states of disappointment and regret emerge and take hold.

The natural disposition dwells, from these, upon what it takes as the mental states of disappointment and regret of the person. The scientific disposition, however, identifies these brain-sign states as communicative modalities derived from the causal states of the brain in relation to its actions, in the case of ‘regret’, one state actual (A), the other hypothetical (C). We *seem* to know we are disappointed or regretful because co-existent as brain-sign is the context of the disappointment and regret – their differential cause, in fact, as categories-of-the-world. But these are not mental states of memories or imaginings for us as a knowing (mental) ‘I’, but rather the play of relevant causal orientations toward the world. We (the ‘I’) exist in brain-sign at each of the brain’s moments as the represented organism in relation to its causal activity.

\*

We see from this that mentalism cannot explain its own categories. For we can regret a lot or a little. We can have shades of regret. We can regret more at one time or another. Some new occurrence can wipe out regret. How is mentalism to explain these? All it can do is suppose that we want something more or less. Or perhaps not regret at one moment while we do at another. Or that something replaces our regret by being an alternative or simply a distraction.<sup>106</sup> In all cases we have no access to a physical causality that would genuinely account for regret’s variance.

Indeed, in mentalism’s explanation, we have to go through (explanatorily) the convolutions of a person’s mental being, which itself remains entirely unfathomable *as* a series of causal states. How is it that making a mistake and becoming aware of an alternative strategy which could have been better makes us regret? These states arrive from nowhere, yet appear (mysteriously) to *be* us.

Brain-sign theory, on the other hand, has direct access to the grounds of ‘regret’s’ variance. Rather than regret wafting into being (without any precipitating causal conditions that we understand in a physical sense), we can reduce the state (as category-of-interaction)

---

<sup>106</sup> A kind of Spinoza-ism. ‘The force of any passion or emotion can neither be hindered nor removed save by a contrary emotion and one stronger in checking the emotion’ (1910; 148).

to its caused conditions. 'Regret' occurs because of the activation of the contrasted relevant networks. 'Regret's' intensity derives from the degree of the networks impact on the potential survival of the organism, as given by its physical activation. Regardless of the accuracy of the survival likelihood (for this cannot be *known* by a brain), the intensity is 'simply' the physical conditions provoking the degree of 'regret'. Variance in that intensity can be directly attributed, at the time of its occurrence, to the level of activation. There is no strangeness of 'regret' subsiding because of some applied drug (a sedative, perhaps). It is not a question of the (purely chemical) drug causing a mental state as regret to be less regret. What is causing the 'regret' (as a neural interpretation) is the difference of the activated neural networks; and one can attribute a lessening of 'regret' in the application of the drug to the deactivation of the networks whose *difference* is the source of 'regret'.

Indeed, we can replace the term 'regret' completely, thus removing any relation to a mental state. The relevant brain-sign state *is* the difference between neural network states A currently activated, though not active, failing to be effective (as B) and contrasted with other neural network states C (activated, but not active) in which the failure is not evident. Now we see what regret actually is: a physical category-of-interaction brain-sign, a condition which mentalism *cannot* specify.

Clearly the variations in the category-of-interaction, being directly attributable to the degree of lack of neural control (for survival) *expressed* from the activation levels of the different neural network states, is a meaningful guide to those differential activations. In principle, though whether or not it would be possible in fact, the brain-sign state's 'strength' condition (what 'we actually feel', as the natural disposition would put it) could be directly related to the extent of the difference, measured by some technological means. This removes completely (in principle) the supposed problem of explaining, in physical terms, the subjectivity of subjective states.

Obviously, if '*Je ne regret rien*', there is a lack of networks B or C which are physically effective, or they are capable of being repressed in strength by other physical means (another topic for brain science). Whether this is beneficial to survival is, perhaps, an open question for biology. What is not the case, however, is that *I* can suppress regret by an act of will, or even by being indifferent to whatever happened. This (mentalist) mode of explanation entirely misconstrues the physical biology.

## 6.12 Development of categories-of-the-world and -interaction. [Approaches 3, 4]

There are three fundamental categories of brain-sign, two of which we have named: of-the-world and of-interaction. There is a third, language – but not language with its mentalist suppositions. We will come to that in the next chapter.

Having discussed the operation of categories-of-the-world and -interaction in the previous two sections, we can now develop them in further detail. However, this is a large topic, and the following must still be regarded as only an introduction.

### 6.12.1 The ontological contrast of brain-sign categories *vs.* mental states

The mentalist supposes that within us are states of our mind or person as which we have an understanding of being in the world. But in fact we *find* we do not know what tomorrow, or even the next moment will bring in terms of our orientation to the world (§4.1.3). One moment we are happy, the next perplexed, thereafter sad. Our orientation to the world is in continual change. This change is not graspable by us in fine detail. As Pirandello said in his diary in 1934, two years before he died: ‘There is someone who is living my life. And I know nothing about him.’ This suggests our ability to know ourselves is impossible because what lives our lives has nothing to do with our own (supposed) experience of it. Why does this great lime tree appear one day as magnificent, and the next forbidding? Why does the sun make us happy, and the rain depressed – sometimes?

The illusion of the natural disposition is that our mental states are as the world in its concreteness. But as various thinkers have asserted, our states are not the same as the world: they are not concrete. While we interpret ourselves as *having* a mental life (which is our concrete existence), we fail to grasp our intrinsic nature.

Both categories-of-the-world and -interaction are neural creations derived from our casual being. The actual lime tree is, on one day, physically related to us in a certain kind of way, on another day another. Those ways define what we are as brain-sign *determined* as the tree’s (so-called) appearance. We do *not* experience the world. The lime tree is a category-of-the-world, and our ‘seeing’ it as magnificent is a category-of-interaction. But (cf. Husserl, §5.2.2), the tree we (supposedly) *see* on different occasions is not the same tree towards which we have different feelings. The tree we ‘see’ *is* different because our causal orientation

to its actuality (which we never see) is different. This does not make the tree we see *not* a category-of-the-world. The point is our ‘feelings’ towards it are not a separate psychological faculty (subjectivity) added to a seeing of actuality (objectivity). Categories-of-the-world and -interaction are interwoven because brain-sign represents precisely the unified orientation of our neural relation to the world *per se*. This unified orientation can change in a moment. We remark to our friend how forbidding the tree seems today; but their cheery response immediately alters our neural relation to the very same tree (actually), and so we now see it differently *and* with a sudden cheerful feeling. ‘Of course, you’re right; I don’t know what’s come over me today.’

This demonstrates a function of brain-sign which is not simply ‘getting me to see things your way’. It facilitates the communication of two physical brains in *their* orientation to the world. Methodologically, the mystery of our states cannot be interpreted *from* those states (taking them as irreducible) because, as Pirandello said, someone (some thing) is living (creating) our lives, and we can know nothing about (we have no access to) him (it).

This is the opposite of Block’s contention (Chapter 1). ‘The starting point for work on consciousness is introspection and we would be foolish to ignore it.’ Block supposes he can make these classifications because consciousness is self-revealing to itself. He does not appreciate that his contention is biologically determined (involving no suppositions or wishes) which cannot escape its physical inevitability. The model ‘Block’ proposes is invalid, for there can be no such thing as self-revelation of the phenomenon to itself (§4.1.2). His contention *as* the phenomenon is a neural construct in the natural disposition (which supposes all can be given), not the scientific disposition, which proposes there is no (conscious) someone to whom anything is given. Block (or rather his brain) never explains how *anything* could be given, nor what giving is.

### **6.12.2 The general contrast of categories-of-the-world and -interaction**

Because there is no ambiguity about whether what is represented in brain-sign really exists or not in the world – nothing in brain-sign is an ontologically separate category beyond a neural sign – the fundamental classifications of brain-sign are greatly simplified (from mentalism) to three. In the following we use words (as we must), but we are describing pre- or non-

linguistic neural states to which those words point. (As said, language is discussed in the next chapter.)

All categories-of-the-world are objects, events or situations towards which our organism has a causal orientation. (In the case of hunting lions, the gazelle is an object, the chase is an event, and the gazelle, other lion and terrain is a situation.) Whether they exist or not is, in principle, beyond our ability to ascertain. Moreover ‘imagined’ categories are still, in brain-sign terms, categories – given the function they perform, i.e. communication about the world. They are a pointing-to what in the world (actual, recalled or imagined) may cause actual, former or imagined action.

Categories-of-interaction specify the *how* of the organism’s reaction to the world. Here are a few examples.

Descriptive adjectives can denote categories-of-the-world. But adjectives of our (organism’s) orientation to the world, still termed descriptive in grammar, are categories-of-interaction. Thus the ‘blue sky’ is a category-of-the-world because both blue and sky are of the world. But the ‘magnificent tree’ designates both the world and the interaction, for ‘magnificent’ denotes a reaction to the tree, not a description of the tree *per se*.

Categories-of-the-world, as the tree, are directly reducible to the brain’s construction of them from the causal conditions which provoke them. Categories-of-interaction, however, require a further kind of reduction.

What does ‘magnificent’ entail? What is signified is the way the organism is *reacting* to the world, and this can be extremely subtle to the point of inexpressibility (because of the complexity of neural function). Yet how the world affects us, which the adjectival word denotes, must be decomposed. ‘Magnificent’, in this case as our *sense* of the tree, refers to scale, perhaps complexity, the vividness of the colour here in June, the fact that it has taken a century to bring the tree to being in the way it is, even the way we have to look up at it. Categories-of-interaction are therefore the signifying sum (or token) of numerous features which are *of* the world, so they are effectively decomposable to the provoking causal conditions (as above concerning regret). I.e. categories-of-interaction have no intrinsic being (as mental states), independent of the world as such (cf. Peirce, §5.2.1). They are not, of course, decomposable to categories-of-the-world (cf. §6.6).

This example exposes the distinction we are making between brain-sign and mentalism. It may be proposed, for example, that the designation of 'blue' is incorrect because the person making it could be colour-blind, and therefore blue depends upon the person and should be a category-of-interaction (in mentalism, the subjective). But this is not so because what is described by 'blue' is the object itself (even if wrong), not a reaction of the organism. On the other hand it may be proposed that everybody would agree the tree is magnificent (in mentalism, the objective); therefore magnificent is a category-of-the-world. But the tree is an object as it is: whether or not everyone would describe it as magnificent does not obviate the fact that 'magnificent' designates an organismic reaction. Indeed, we have said that categories-of-interaction are decomposable to what in the world has provoked them. The distinction is that the brain *creates* a (decomposable) category-of-interaction precisely to communicate reaction.

A similar designation occurs with verbs and adverbs. A 'seen' act of running, as brain-sign, is a category-of-the-world. And if the person is 'seen' running swiftly, that too is a category-of-the-world. But running beautifully introduces a category-of-interaction.

The how of the running indicates the need for decomposition. 'Beauty' has to do with shape, ease of movement, regulated composure, etc. These decomposed categories *are* of the world, and thus can be directly reduced as brain-sign, while their aggregation as the *sense* of beauty, is a composite neural (brain-sign) shorthand.

Can dogs see 'the beautiful'? Clearly they cannot, and neither can we. But that does not mean the brain-sign of dogs does not assemble in such a way that shape, ease of movement, regulated composure are not signified as, as it were, an less developed equivalent of what, in humans, is signified as the sense of beauty. After all, 'beauty' is causality engaged in a way so signified, and it is not improbable a canine equivalent exists. Obviously they do not express the concept.

But like 'magnificent', 'beauty' is a category-of-interaction applied with many different kinds of categories-of-the-world. Categories-of-interaction do not have the precision of categories-of-the-world, and this reflects the way they are made (or decomposable).

\*

But as said, categories-of-interaction not only replace the natural disposition's (1) emotions or (2) aesthetics. They function with categories-of-the-world as e.g. our (3) use of them.

For example, a fork is usable *per se* (3), and this relates to its form (allied to aesthetic notions (2)). Our ability to use the fork registers as an emotion, or feeling (1). What unites these categories-of-interaction *in toto* with the category-of-the-world is their signification in the unit of the communicated. The infant's *sense* of bodily use of the fork *qua* fork (§6.6) is conjoined with their 'pleasure' (as 'feeling') in being able to use it. This is *signified* as the organism's (new) dexterous and socially accepted engagement with the fork. The adult has long since 'forgotten' the 'pleasure' of using a fork (cf. habit, §3.3.1), but displeasure immediately occurs if the design (the aesthetics) is poor (cf. Heidegger, §5.2.2).

In like manner we can describe the lions. The gazelle is a category-of-the-world. Its appearance as 'eatable' is a category-of-interaction (3), along with its 'edibility', or appearance (2). The 'excitement' of the natural disposition (§5.5) is also a category-of-interaction (1) in the chase. The extent to which lion brains have these states is conjectural, which may or may not be open to experimental confirmation. However the scenario offers the possibility of unproblematic evolutionary continuity.

### 6.12.3 What mental state terms are as brain-sign categories

Nouns directed at mental states, like sight or hate (or regret), signify behaviour invoking conditions. In brain-sign terms however, they are ontologically displaced. Sight does not exist. Therefore 'sight' is a *displaced* state, exactly as (unscientifically) we say the sun rises in the morning. 'Sight' points to a category-of-the-world, but, in natural disposition terms, it is an imaginary state.

Hate is complex. The word relates to brain-sign states of varying sense and intensity signifying behaviour (actual or potential) – e.g. violence or avoidance for the category-of-the-world. It is a category-of-interaction.

Does love designate a particular state? Is the brain-sign *sense* to which the word 'love' points a precise state? Clearly the answer is: no. Probably even more than hate, different behavioural kinds evoke conditions causing the word 'love' to be associated with them. The various *senses* cannot be regarded as identical, because what is involved in behaviour can be quite different. Therefore trying to define what love really is is futile, for 'love' is not a category-of-the-world. But it does indicate a behavioural opposite to hate, *viz.* the sense of a kind of unison or union with the object of a category-of-the-world – the

neurophysiological condition associated with keeping the object, actual or not, present. (Loss, so-called, is thus the signified absence of the object, event or situation.)

\*

In the natural disposition we may say we are anxious or nervous about an event because we do not know how it will turn out. We say our anxiety results from the unknown which will clarify itself once the event occurs. We may be irritated with ourselves for being nervous: ‘Why am I like this? Why can’t I control myself? I know what the reason is, but why can’t I *not* worry about it?’ Here is the extant confusion of the world with our inner states. How could a condition of the world (a physical event) cause anxiety (a mental state)?

From the scientific disposition the analysis is different. What causes our ‘anxiety’, a category-of-interaction, is the brain’s inability to effect a precise outcome because it cannot act to do so *now* (cf. §6.10). In reductive terms, because the brain cannot act, it cannot control, and thus ‘anxiety’ is a neural invention signifying lack of control.

As brain-sign we may ‘hope’ for the best, or ‘imagine’ the worst. Both these categories-of-interaction (best and worst) are interpretations by the brain of its own activated but not active causal states (the networks signified by categories-of-the-world). But what defines *us* at this moment is the *causal* state of the brain in its impotence because, while it can be relevantly activated, it cannot act or react actually. ‘We’ (as brain-sign) cannot alter this. Our inevitable suffering is the communicated state of ‘anxiety’.

There may be a benefit in the scientific disposition because the neural mechanics are defined, rather than being an undefined ‘mental’ turmoil. There may consequently be a lessened ‘anxiety’. For supposing we have an unaccounted for but causal mental life is like wondering (without aerodynamics) how airplanes stay up in the sky. (We will address this more extensively in Chapter 9.)

#### **6.12.4 Simulation theory and theory theory**

The problematic nature of sympathy and empathy have been posed in philosophy for centuries. How can we associate ourselves with or react to what is going on in others? But of course, while sympathy and empathy are understood to be mental states, they lack any cogent (i.e. reductive) explanation. Two recent competing accounts of them are: *simulation theory*,

i.e. the subject undergoes equivalent states to the person for whom sympathy or empathy is felt; and *theory theory*, i.e. the subject has a psychological understanding of the state of the other person's condition (cf. e.g. Davies & Stone, 2005). Neither account explains the situation because mental states are inexplicable; but brain-sign theory, in a way, unites them *via* the unit of the communicated.

It is likely that our 'grasp' of another's condition results from an equivalency in us of their physical orientation to the world, which is brain-signed. For example, if they are physically impaired and in pain (categories-of-the-world), their physical orientation to the world is degraded. This is a biologically reducible condition of the organism in terms of its capacity for survival and reproduction. Our organism can replicate this condition without its consequent actuality *via* neural reformulation (§6.10). We might term this simulation. Resulting from the simulation is the brain-sign of 'psychological' equivalency, i.e. the 'distress' (category-of-interaction), with the *additional sense* that, for us, it is *not* actual – because it is not. The additional senses beyond the distress are the brain's invention of sympathy, and to a greater degree of neural equivalency, empathy (as categories-of-interaction), as the natural disposition terms them. Theory theory wrongly considers these psychological understanding, for that, of course, is undefined physically/ontologically.

'Sympathy' and 'empathy' as brain-sign are examples of *reciprocity* in the unit of the communicated (by contrast with the lion scenario, where there is a common target of the hunt). The unit of the communicated is one physical structure of the two parties, possibly leading to action like help and support. Brain-sign does not cause the actions, but signifies the neural reformulation (as 'sympathy and 'empathy') which results in a physical way (help and support), a socially programmed neural response. The receiving party may then have evoked (as category-of-interaction) what the natural disposition would term gratitude.

The reciprocal parties are bound in their physical actions in a way inaccessible to brain-sign, and the reciprocal brain-signs are impenetrable to each other. There is, thus, both a sense of the 'why' of the actions (the brain's self-explanation), but also a sense of an invisible but unbreachable boundary between them, i.e. their absolute difference and condition which cannot, as is (physically), be represented in brain-sign. Thus the haunting sense of individual isolation. Obviously this does not arise from the difference in fact; rather the oscillation of each individual between the activated networks (the sufferer), and the active but not activated networks of 'simulation' (the responder, as in §6.11).

### 6.12.5 Brain-sign eliminates the subjective/objective distinction: Consider pain

Now someone might still say: ‘But all you have done is propose categories more properly understood as objective versus subjective. Categories-of-the-world are *of* the world, whereas categories-of-interaction depend upon the individual.’

Were such a statement made, it would be necessary to point out again that categories-of-the-world and -interaction do not function under the rubric of mentality. Their function is purely to signify the cause of neural causality.

Indeed, since the causal connection of the organism to the world is its physical states, for survival and reproduction rather than knowledge, ‘subjective *vs.* objective’, or the ‘first person *vs.* the third person point of view’, are explanatorily useless. The ‘I’ is not the foundation of brain-sign, and therefore there is no mental subject; and brain-sign results from the causal status of the brain which cannot be objective. Doubtless in ordinary speech we will continue to use such terms. The pedagogic issue is that they are used in (supposedly) instructive texts of philosophy and science without justification beyond the custom that attaches to them.

Moreover, subjective categories can include the blue of a blue sky. Brain-sign theory does not accept this because, in terms of neural manufacture, blueness is as much a part of the brain-sign *world* as the sky (insofar as one can refer to the sky independently of its colour).

More generally, ‘feelings’ and ‘sensations’ do not necessarily correspond to categories-of-interaction. Pain, for example, is not a category-of-interaction; it is a category-of-the-world because it identifies a physical event (actual or otherwise) in the world of an organism. The related category-of-interaction is the *sense* of distress at whatever level obtains. Even pain from a phantom limb is a category-of-the-world, for what is signified is the non-existent injury of the non-existent limb. This clarifies pain in the physical universe, of which Patrick Wall said:

‘We have to clear our minds of the idea of a rigid signalling system. It is simply not true that a particular injury generates a fixed amount of pain, and that feeling means injury.... Philosophers [have] generated a confused mess [in] construct[ing] idealised

“rational” sensory systems in happy ignorance of the working of living organisms’ (1999; 1) (slightly reconstructed).

We would say rather that mentalism is a ‘confused mess’.<sup>107</sup> The methodological effectiveness of the brain-sign account of pain can be seen from this footnote of Frith.

‘The mental experience of pain is associated with activity in the anterior cingulated 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; 150f).

If we remove the references to feeling and emotion, the situation is more effectively explained by replacing pain itself by a category-of-the-world, and our reaction as a category-of-interaction. For how could pain be *pain* (as felt) if there were no emotional reaction to it? (Frith, of course, has no alternative model to explain the situation.) The brain-sign account demonstrates that before cingulotomy, what is associated with the category-of-the-world (‘pain’) is a negative category-of-interaction. Once excision has taken place the category-of-the-world remains, but the category-of-interaction has changed to ‘indifference’. Thus neuroscience is offered a more physically discriminatory account for investigation.

#### **6.12.6 Brain-sign terminology will not replace mental terminology generally**

The benefit of the brain-sign structure is its bio-scientific aptness, and so its path to reduction. It is probable, however, that normal life could not support this (limitless) analysis in conversation. Indeed, the vast complexity of causal neural conditions are interpreted by the brain itself into signifying states of a *very* limited range or content (comparatively), which is how communication *can* take place between organisms.

#### **6.12.7 The question of terminology**

We have continued to associate words like ‘regret’ and ‘anxiety’ with brain-sign states, though we have said they do not denote mental actualities and we put them in inverted

---

<sup>107</sup> Howard Fields and Donald Price conventionally state: ‘The experience [of pain] is personal.’ And “‘Pain” refers to a subjective experience’ (1994; 452).

commas. We refer to the tree we ‘see’, even though we do not see the tree. Is it right to use these words at all? (We showed at §6.11 how the words could be removed completely.) After all, when the phlogiston model was replaced by that of oxygen (in molecular chemistry), use of the word phlogiston disappeared.

Two things can be said. First recall Fodor again (Chapter 1). ‘If commonsense intentional psychology really were to collapse, that would be, beyond comparison, the greatest intellectual catastrophe in the history of our species.’

Fodor confuses the communication the words of commonsense psychology effect with what they denote. Indeed, he wishes words to point to actualities – beliefs, desires, etc. – even though he is aware of the problems involved in the causal functioning of words (§3.1.4). By contrast, we have demonstrated that words can point to something that commonsense psychology (or the natural disposition) supposes, e.g. sight, but without effecting its natural disposition function. By so doing we have changed the context from psychology to neuroscience. (However, until we arrive at how language works in Chapter 7, we lack the means fully to explain words.)

Equivalently, when the *solar* system was identified, planets retained their geocentric names, though they were no longer the entities previously supposed, nor conformed to the geocentric model in behaviour. Thus there is a precedent for retaining the names provisionally.

The deeper point is, however, that there are four tasks involved in fully replacing mentalist terms. (1) We must create a vocabulary for the biological conditions and structures of neural causality in relation to the world *per se*. (2) We must identify how, in the brain, these are carried out, at a global, thematic and micro level. (3) We must analyse more specifically what the content of brain-sign states are in their role of signifying causal conditions. And (4) we must identify how brain-sign states exist in the brain. (The first two, of course, are in progress, although as yet without the notion of brain-sign theory. Aspects of these are discussed philosophically in Bickle 2009.)

Since brain science, as here outlined, is still in its infancy, extensive work on a more appropriate vocabulary must be inevitably limited. This vast subject is to be approached in steps. Our first step here is to introduce the concepts of brain-sign and replace consciousness with it. (We will refer briefly to this again in Chapter 10.)

### 6.13 Music. [Approaches 2, 3, 4]

An analysis of music reveals what brain-sign theory says about *us* as organisms which is quite different from what we naturally suppose. For the natural disposition, music is a joy for the individual person: it can evoke passion and intellectual gratification, excitement and despair. But in terms of the scientific disposition music functions quite differently.

Before addressing music *per se*, we consider Steven Pinker's book titled *How the Mind Works* (1997).<sup>108</sup> In the blurb, the following occurs.

'This book is about the human brain, though not directly about neurons and hormones; that is because the mind is not the brain, but what the brain does.'

Taken from a passage early in the text, this passage is astonishing because it is impossible to grasp what it means. However, an assessment of it is revealing.<sup>109</sup>

The brain is certainly neurons and hormones (amongst other things), but how can it be that the book 'is about the brain' but 'the mind is not the brain'? If the mind is not the brain, then what does the book purport to be about in claiming it is about the brain? At a pinch one might suppose that some kind of Wittgensteinian trope is proposed; that two different kinds of discourse are in play, one in brain terms (neurons and hormones), and another in mind terms (perceptions and thoughts). But this is evidently not the case because Pinker also says that the mind is 'what the brain does'. How are we to understand that? Are we to suppose the brain is like a motor car, but what a motor car does is move? Therefore the mind is to be thought of, equivalently, as the car is in its speed and direction? But even if such a notion were plausible, speed and direction belong to scientific categories: of motion-dynamics and topography. Mind terms have no underpinning scientific classification, so we cannot simply assume that perceptions and thoughts are an alternative scientific category to which 'what the brain does' can be attributed. (We can talk about brain-sign and the virtual array because signs are physical.)

---

<sup>108</sup> Two more recent books, by Daniel Levitin (2006) and Oliver Sacks (2007), refer to Pinker. Sacks actually quotes Pinker's position which we analyse in greater depth. Neither make the necessary ontological distinctions explored here.

<sup>109</sup> Though it is likely Pinker has in mind the computer analogy of hardware and software, which is inapposite.

Being an evolutionary psychologist, Pinker is committed to the idea there is a *psyche*, and the terms of its domain, for example perception and thought, are *de rigueur*. Unfortunately, this is not scientifically valid, and Pinker never explains why it should be. The kind of difficulty he gets into is instructive. Here he is on music.

‘Music is an enigma.... As far as biological cause and effect are concerned, music is useless.... Music could vanish from our species and the rest of our lifestyle would be virtually unchanged’ (1997; 528). And then: ‘I suspect that music is auditory cheesecake’ (535).

According to Pinker music, by which one must assume he means what we hear, should be a biological category in the terms he expresses it: i.e. for evolutionary validity (for survival and reproduction) it must entail cause and effect. Since evidently it does not, hearing music is superfluous and is to be dismissed as ‘cheesecake’. What Pinker means by cheesecake in scientific terms is not explained – but it is clearly pejorative.

Hearing music, of course, must be regarded as a mind state, and since Pinker says that his theories do not involve the brain in neural terms – even though cause and effect biological categories can only apply to neural terms because it is they that are physical – his claims concerning heard music (i.e. mind states, not brain states) are already suspect. Soon the structure he assumes begins to collapse. He says:

‘Auditory perception is inverse acoustics: the input is a sound wave, the output a specification of the sound-makers [e.g. musicians] in the world that gave rise to it’ (*ibid.*).

We see immediately what is wrong. *Sound* waves do not exist: there are compression waves, which are purely physical. Thus the ‘output’, what we hear (sound, notionally), is not a specification of the ‘sound-makers in the world that gave rise to it’, because the two are entirely different – i.e. there are no *sound*-makers: what is *made* by the so-called ‘sound-makers’ are compression waves which no one can hear. ‘Auditory perception’, therefore, does not *correspond* to ‘input’, i.e. compression waves. Pinker muddles the world with states of the brain phenomenon.

In claiming (without questioning it) that mental states exist, Pinker supposes what we hear is cheesecake because music is enjoyment, rather than serving a cause-effect biological

function. But if there are no mental states, then the rationale of evolutionary psychology is *de facto* invalid. He continues:

‘One of the brain’s tricks as it identifies the sound-makers in the world is to pay attention to harmonic relations. The inner ear dissects a blare into its component frequencies, and the brain glues some of the components back together and perceives them as a complex tone.... Presumably the brain glues them together to make our perception of sound reflect reality’ (*ibid.*).

Having made the mistake of sound-makers – i.e. supposing that sound exists in the world – Pinker then proposes the brain ‘pay[s] attention to harmonic relations’. This encourages the vision of the brain working out what harmonic relations are so that it can make of them a ‘complex tone’. (Indeed, we are told that the brain can ‘perceive’ a complex tone, which is absurd.) The brain does no such thing. The brain is the locale of mindless operation. The brain receives physical input transduced *in physical type* from the ear, which at the ear is certainly not a ‘blare’ since it is not sound, which ‘blare’ suggests. Therefore the inner ear does not ‘sort out frequencies’ which the ‘brain glues...together to make our perception of sound reflect reality’. ‘Reality’, Pinker’s word, is an unidentified category and never addressed. Note that in this sentence, an identity is presumed concerning the word ‘our’ (‘our perception’) and the ‘brain’ (the brain ‘perceives...a complex tone’), the very identity that Pinker says in the blurb he is not going to discuss.

Since we (as organism) exist in a physical universe, our only reasonable assumption is that *reality* (insofar as this is a meaningful term) must be physicality. Both the creators and hearers of music exist as physical objects. Therefore what passes between them must be the physicality of music, i.e. structured compression waves. What is impacted is their physical brains. Thus the significance of music is its *neural* impact. What is ‘heard’ in both creators and recipients of music must be what is shared between them as organisms; it is a *communicative* modus, not a mental (or psychic) state.

There are two things to be said. The first is about Pinker’s approach. Though he denies concern with the brain, he proceeds (inevitably) to introduce brain function. As is then obvious, his discourse is too ontologically confused to be intelligible.

More relevantly *vis-à-vis* our nature: When, in the natural disposition, we consider ourselves enjoyers of music, we take for granted this is because we do *hear* music – the

egocentric orientation as which we are created: music is present to us in our hearing. But what is significant about music (in the scientific disposition) is what we cannot ‘hear’, even though it is we (as organism) who are undergoing it, *viz.* the physical effect of compression waves on our brains. Music is *physical* communication with others. But the ‘we’ in the sentence is not the psychological subject: it is us as physical organism.

It might seem extraordinary that our pleasure in music has nothing to do with *our* having it. That even alone in a room ‘listening’ to a CD, our so-called experience is our brain communicating. But once this is grasped, our biological existence (bio-centrism), and our relation to other organisms physically, will also be ‘grasped’.

Is music cheesecake? With no biological cause and effect? Pinker continues:

‘[Darwin] suggested, not too plausibly, that human music grew out of our ancestors’ mating calls. But his suggestion may make sense if it is broadened to include all emotional calls. Whimpering, whining, weeping, moaning’ etc. (536-537).

This is an attempted reduction, to originate the sophistication of music in a more primitive biological conduct. Association with ‘emotional calls’ would have music derived from a cause-effect modality. Indeed, emphasising the communicative aspect could be seen as a move towards brain-sign theory. But (being cheesecake) music does *not* function with that communicative purpose, Pinker must say. So the move would be stillborn.

In fact the move cannot work because it is the wrong reduction. Psychological states cannot be reduced to more primitive psychological states because we have no reductive base for psychological states *per se*. And since Pinker has ruled out discussing the brain (at least he says he has) where reduction *would* take place, reduction is impossible.

Churchland rightly says that philosophical explanation must ‘go into the brain’ (2007; 238). So, in bio-neurophysiological terms, the impact of music is to alter our brains. In brain-sign content we seem to find passion and excitement, etc. But transduced *compression waves* have a different effect. They improve our concentration (cf. Sridharan *et al.*: 2007); they calm our ‘irritation’ or ‘anger’; in war, they drive soldiers forward. Thus the physical impact on the brain alters our behavioural functioning. Soldiers ‘hear’ (roughly) the same stirring marches. But what they hear serves neural *communication*, not the drive forward.

It might seem extraordinary to propose that a function of music is *merely* to calm us. But that does not capture what is involved. While calming might occur, the impact of music as compression waves is *neural* organisation, which is a highly complex activity to be understood in neural terms. ‘Hearing’ is a transformed précis by the brain of that neural organisation *for* communication: a much simpler activity, however important it might seem to us (in the natural disposition). In other words, those involved in creating music do not understand what they are doing actually (until the reduction is spelled out), and neither do we as ‘listeners’.

\*

Because music is ambiguous between category-of-the-world and -interaction, its nature has seemed hard to identify. While it coincides with no object in the world, it seems to evoke such because, as category-of-interaction (e.g. ‘excitement’), there is a sense that ‘out there’ substantialities should exist with which it coincides. Indeed, music is shared as being *in* the world, but is reacted to in a specific way by each of us. Schopenhauer said:

‘The inexplicable depth of all music, by virtue of which it floats past us as a paradise quite familiar and yet eternally remote, and is so easy to understand and yet so inexplicable, is due to the fact that it reproduces all the emotions of our innermost being, but entirely without reality and remote from its pain’ (1969; 264).

Schopenhauer captures the apparent ambiguity of music as without worldly substantialities (‘without reality’) yet hinting at them (‘a familiar paradise’); and the ‘reproduc[tion] of all the emotions of our innermost being’ emphasises the direct action of compression waves on, amongst other locations, those delivering (after neural self-interpretation) what are referred to as emotions – though the compression waves impact more neural function than signified by ‘emotions’.

Yet, despite the apparent ambiguity of music, we *can* clarify beyond Pinker’s claim about cheesecakiness.

It may be (as Pinker proposes) that without objects in the world there seems nothing to apply cause and effect to in music. However compression waves *do* function precisely in cause and effect – in terms of neural organisation. It is the neural brain *itself* that causes our organism to engage with compression waves (which we ‘hear’ as music), and it does so in its *causal* neurobiological functioning, which certainly cannot be regarded as cheesecake. To

‘understand’ the biological impact of music requires an ‘understanding’ of its neural organisational function.

Pinker’s analysis fails in a similar way to Bennett and Hacker with perception. Because he assumes he already knows what mental (or psychological) states are, he has not developed the required explanatory structure. Indeed, he deliberately does not consider the brain in neural terms because (by some unexplained formula) the mind is what the brain does, and only the mind matters. Pinker’s analysis is the unfortunate but illuminating consequence of the natural disposition.

Though one would wish far more on this topic, since we wish to ‘grasp’ what the relation between neural organisation and our state as brain-sign *is*, we must leave it here.

#### **6.14 What do deficit and illusion show? [Approaches 1, 2, 3, 4]**

No direct experimental work on brain-sign yet exists. In the preceding text, practical examples from the literature, in their reinterpretation, have demonstrated the explanatory benefit for neuroscience of brain-sign over mentalism. Since brain-sign theory claims to be a scientific theory, they are essential. We now add briefly to those examples.

##### **6.14.1 Deficit**

When the work of Milner and Goodale on deficit was first introduced (§3.3.3), we drew attention to one particular sentence about tasks Dee could not perform because she had not the communication skills. ‘Dee could do none of these things – not because she couldn’t communicate, but because she had nothing visual to communicate. She had no conscious...visual experience...to share with us.’

We did not engage with it, but now we see its significance. Although Milner and Goodale *identify* communication, they understand it as a derivation from consciousness as a knowledge-type function. However they also say:

‘The most amazing thing about Dee is that she is able to use visual properties of objects such as their orientation, size and shape, to guide a range of skilled actions – despite having no conscious awareness of those same visual properties.’

It should now be clear that consciousness (i.e. vision, as in ‘visual properties’) is irrelevant. Her deficiency was in the structure of the human super-organism. While Dee could act, she could not be part of the unit of the communicated. Milner and Goodale’s account is scientifically opaque because they neither explain consciousness nor how communication works. Consciousness obscures what science requires. Brain-sign theory explains the findings of Milner and Goodale with a genuine physical model, whilst still preserving, indeed enhancing their significance.

(The same kind of conclusions can be drawn concerning *blindsight* in the work of Lawrence Weiskrantz, but it will not be engaged here.)<sup>110</sup>

Other examples than pain demonstrate the benefit resulting from the distinction between categories-of-the-world and -interaction. Deficits occur which cause the inability of the subject to identify the use-function of categories-of-the-world, i.e. their associated categories-of-interaction, or attaching the wrong use-function to categories-of-the-world. Such deficits, and the difficulties subjects have with them, are well documented in the literature, and have been popularly described by Oliver Sacks (in e.g. *The Man Who Mistook His Wife for a Hat*, 1985).

### 6.14.2 Illusion

Psychological texts are filled with examples of illusions. But what do they tell us? For example, Frith says:

‘Epilepsy affects about 1 person in every 200.... Just before a seizure occurs, many sufferers start to have a strange experience known as an “aura.”... For one person it may be the smell of burning rubber. For another person it may be a buzzing noise’ (2007; 31).

Frith claims the brain ‘makes up the mind’. But how should this illusion be interpreted? If the brain creates illusions for the epileptic, for no burning rubber or buzzing noise are in the vicinity, we must conclude that ‘the strange experience’ is the brain *self-*

---

<sup>110</sup> In e.g. Weiskrantz (1997). Milner and Goodale make an analogy between their own work and that of Weiskrantz.

interpreting. Causal neural activity exists, the physical seizure, but its prior ‘aura’ experience is a neural interpretation of prior neural activity.

‘Smelling’ actual burning rubber or ‘hearing’ actual buzzing as brain-sign must be the brain interpreting the very same areas that cause the illusion precursive to an epileptic fit. In the actual case, the function of ‘smell’ or ‘hearing’ is to communicate with another. But *exactly* the same function exists with epilepsy. There is no illusion of *experience* (for nothing is ever experienced), but a neural contrivance for communication in the physical world – the so-called ‘aura’. Of course, the brain does not do this deliberately, as it were; the effect is contingent upon the particular circumstances of neural behaviour in epilepsy.

But why are there illusions in the seeing of a *normal* person? Frith (conventionally) puts this down to the visual system in its genetic design.

‘Our brain is hard-wired during the first few months of life as a result of our visual experiences. There are certain facts about the world that do change very little, and therefore, become strong prior hypotheses’ (128).

Thus (as we have discussed) since so-called perception (as brain-sign) is a constant re-occurrence, psychologists can devise experiments causing us to ‘see’ the actual world wrongly because we are *not* seeing but imposing upon the world the re-occurrence as which our ‘perception’ is constructed (by genetic design).

However, Frith’s interpretation of illusions depends upon the notion that we are conscious, even though it is evident that not seeing the world *as is* would seem to imply, by definition, we are *not* conscious. Recall Bennett and Hacker:

“‘To become conscious of (or that)’ and ‘to be conscious of (or that)’ are factive, or existence implying – that is, one cannot logically become conscious or be conscious of something that is not the case or is not there.’

Underlying all this, as Milner and Goodale point out, is the fact that there is not one visual system but many, and for different functions. Vision has been aggregated in evolution over millions of years, and new features are added to previous structures, potentially causing anomalies in appropriate circumstances.

What Milner and Goodale are really referring to here is vision as a physical mechanism. The notion of consciousness as a *method* of seeing (i.e. as perception) is quite another matter, and it is important to unpick what has been uncritically presupposed.

For example, Milner and Goodale report an experiment (the Tichener circles illusion) in which the size of a poker chip to be picked up is ‘disguised’ by surrounding chips of a different size, giving the (‘experienced’) illusion it is a different size from what it is. Yet the experimenters found that, despite the visual illusion, the experimentees formed their grip size to the actual size – i.e. they picked up the chip accurately while still subject to the illusion (2006; 167-168).

But there is much more to this, and more profoundly so for science, than the mere differentiation of motor control vision and conscious vision – as it may first appear. If, as we have said, the only way a ‘seeing’ can exist (as brain-sign) is by the brain’s *self*-interpretation, then the experimentees illusion in the ‘seeing’ does not result from: (1) What is actually in the world – because what is there *is* what is there. Nor does it arise from: (2) The brain’s self-interpretation of the motor mechanisms that caused the accurate grip size – because the grip size was indeed accurate.

Yet the ‘seeing’ *must* be a neural self-interpretation (as brain-sign) – there is nothing else it could be. Thus what *causes* the illusion (the reason it happens at all) must be a neural self-interpretation of the visual system as an anomalous whole (or as much as is entailed). And this has nothing to do with *consciousness* as a casual function (*pace* Frith).

The implication (which supports brain-sign theory) is that illusory seeing, like all so-called seeing, *is not for the purpose of action success by the individual organism*. The error causing the illusion *enters* with the neural self-interpretation subsequent to motor control. Thus neural self-interpretation as brain-sign (rather than seeing) is an *additional* function of the brain for communication, *not* for action. (See diagram at §6.5.)

## **6.15 Conclusion**

This chapter has addressed many topics in brain-sign theory. The complexity and extent of the theory, and its implications, will have become apparent.

In the coming chapters further aspects of the theory, and its import, will be introduced.

# The Theory of Brain-Sign

## 7.0 Language

In the natural disposition language, like consciousness, is a mystery. And like consciousness, too, there is no commonly accepted explanation of it.<sup>111</sup> So despite its overwhelming significance, it is an unknown. Therefore the being of a human being is essentially unknown (cf. §3.3.4).

The third category of brain-sign is language. But since a scientific account of language is located within the framework of brain-sign theory, it is not as understood by the natural disposition.

This chapter is in three stages. First we will establish the context of brain-sign theory on language (§7.1.x). Then, by contrast, we critique various other approaches in the literature (§7.2.x). Finally we specify how brain-sign theory defines language (§7.3 onwards).

### 7.1 Preliminary considerations

Texts on language customarily begin without considering whether mental states exist. That they exist, of course, is crucial to the natural disposition's account of language. It will serve us well in the scientific disposition, therefore, to begin by standing back from mentalist assumptions and reflect upon the territory, employing the material already covered.

#### 7.1.1 The general situation

When someone speaks to us we suppose we hear their voice. But that is not the case. What we hear is a manufactured state as brain-sign. Indeed, we do not hear *their* voice. There is no voice. And we do not hear.

Why do we 'suppose' we hear a voice? Because as brain-sign, we are the neural communication mechanism. In other words, we are not a (conscious) subject for whom a

---

<sup>111</sup> E.g. Paul Horwich begins his book *Reflections on Meaning* with these words: 'Each expression of language surely *means* something – there is some fact as to what it means; but the nature of such facts is notoriously obscure and controversial' (2005; 1).

voice is heard as which we have knowledge of it. When the 'voice' appears it appears as us as brain-sign.

Of course, there may be other features of us as brain-sign co-present. The speaker may be in our view. He may be set in a landscape. He may look particularly dapper today. The totality of our state as brain-sign has the sense of the whole world present to us, including the voice of the speaker. But presence to us, that essential notion for philosophy, is an error of the natural disposition. Nothing is present to us: we (as brain-sign in its totality) *are* present as the neural communication mechanism with, in this case, the speaker. The speaker is a reciprocal brain-sign state, allowing brain communication to take place.

To be complete. In the exchange of language with the speaker, we also suppose we understand the context of the exchange. But as a biological construct, no conscious subject *additionally* knows or understands that we are listening to the speaker as well as doing so. That *sense* is still part of the communication structure. That we need not communicate what we 'suppose' we understand along with our 'hearing' does not mean it *is* knowledge of the speaking process and for us alone (in our minds). Nothing in brain-sign is for us alone, even though it need not be communicated explicitly.

### 7.1.2 The relevance of Saussure

Most other languages than our own are incomprehensible. When we hear two foreigners talking swiftly in their language, we are completely at a loss as to what they are saying. Yet they are using language which *they* seem to understand.

When the majority of us 'hear' the sea we seem to know it is the sea. But 'hearing' the word 'sea' and 'hearing' the sea are different things. For the word 'sea' in its form in other languages means nothing to us who do not know those languages. 'Hearing' the word 'sea' in our language may *seem* to cause us to 'picture' the sea or 'hear' the sea. In another language no such thing happens.

So what is the word 'sea'? It is entirely arbitrary. As Saussure says:

'The bond between the signifier [the heard sound] and the signified [the concept so signified] is arbitrary. Since I mean by the sign the whole that results from the

associating of the signifier and the signified, I can simply say: *the linguistic sign is arbitrary*' (1970; 45).

He then correctly states:

'No one disputes the principle of the arbitrary nature of the sign, but it is often easier to discover a truth than to assign to it its proper place.'

The difference between our account and Saussure's is that we do not suppose the linguistic sign, for example the word we 'hear' and its association with the concept or image, is a mental state (which would imply we understand what the word 'sea' refers to). The 'heard' word 'sea' is not linked to the sea as concept or image *for us*. Indeed, it is precisely because it does not so exist that we can function as brain-sign as physical communication with others.

Of course we can say: 'The word "sea" is what, in English, refers to the sea. Since I know this, I can use the word "sea" appropriately and in context.' And in our head we may *seem* to find that the word 'sea' does evoke concepts or images, and *vice versa* when we 'see' the sea. But the occurrence of the word 'sea' as 'heard' and the associated images is what we already are (as brain-sign), not what we know together (as mental subject). One does not evoke the other, for what evokes both is the physical process *as which* they are associated, which we cannot know as brain-sign.

We can see this is the case because while the 'sound' of the sea identifies the sea *per se*, 'hearing' a foreign word for sea evokes nothing. Saying that for us the 'heard' word 'sea' evokes the sea (as concept or image) does not explain how something entirely arbitrary (the 'heard' word) can refer to something particular (the sea). For neither what we 'hear' as the word 'sea' nor the 'sound' of the sea itself are in the world *per se*. What has become associated are two entirely different impacts of compression waves on our ears (those of the sea and the spoken word), *thence*, in terms of neural interpretation as brain-sign, as the 'heard' word and the 'sound' of the sea, which are neither heard nor are a sound. Saussure says:

'Even if people were more conscious of language than they are, they would still not know how to discuss it. The reason is simply that any subject in order to be discussed must have a reasonable base.... But language is a system of arbitrary signs and lacks the necessary base, the solid ground for discussion' (1970; 49).

Without appreciating the import of his words, Saussure states precisely why language must be interpreted as brain-sign and not ‘conscious’(!) mental states, not least to be discussable. The significance of the arbitrariness of language lies not in what is heard, but in the physical domain as which it is created (and thence ‘heard’). Our organism responds physically/causally to the world *via* the compression waves from the sea *and* the spoken word. ‘Hearing’ (as brain-sign) is the mode by which physical organisms can be in communication about the cause of their actual or potential actions (as for all brain-sign). What we ‘hear’ (as brain-sign) explains our neural causal orientation to the sea *and* the spoken word for others in the unit of the communicated. Thus our causal orientation to both word and sea are linked, though one is arbitrary and the other not.

Our ‘concepts’ or ‘images’ of the sea (evoked by the compression waves of the word ‘sea’) may be different between us. But the adequacy of their overlap is what constitutes neural communication.

## **7.2 The literature on language**

Having initially positioned our functional/structural account of language, we will now address the literature to expose, generally, its scientific inadequacy (i.e. physical implausibility), before returning to our account to express it more fully.

### **7.2.1 Thinking and communication**

There are two supposed conditions of language. The first, that it is the means by which we think; the second, that it is a method of communication. These are obviously entirely different in human ontology. For whereas language (whatever it is) *is* a means of communication, since it passes (apparently causally) between organisms, thinking is a mentalist *supposition* with language as its medium.

Physical actualities make the difference. Speaking causes compression waves to pass between organisms, and writing is the means for electromagnetic radiation to enable textual interaction. But (mentalist) thinking takes place within the head and is not apparent to others, and is therefore not communicative.

Thinking, of course, has been of the utmost significance for philosophers, and to human self-perception generally. It is cognition *par excellence*. Units of knowledge linearly codified. Descartes' formula for the human being is where human modernity *as mind* begins. But does thinking exist?

\*

The physically significant difference between thinking and communication has not been generally appreciated. In his book *Kinds of Minds*, Dennett has a chapter titled 'The Creation of Thinking'. Characteristically Dennett mixes up thinking with communication. Here is a representative passage.

'We human beings have the capacity for swift, insightful learning – learning that does not depend on laborious training but is ours as soon as we contemplate a suitable symbolic representation of the knowledge. When psychologists devise a new experimental setup or paradigm in which to test such nonhuman subjects as rats or cats or monkeys or dolphins, they have to devote dozens or even hundreds of hours to training each subject on the new tasks. Human subjects, however, can usually be told what is desired of them. After a brief question-and-answer session and a few minutes practice, we human subjects will typically be as competent in the new environment as any agent ever could be. Of course, we do have to *understand* the representations that are presented to us in these tests, and that's where the transition from ABC learning to our kind of learning is shrouded in fog' (1996; 175-176).

This passage proposes two different things. Language allows humans to tell each other what to do (he seems to point to both reading and hearing), and thence be able to do it speedily. That is communication. But then Dennett introduces the problem. Thinking, specifically here, *understanding*. What is it to understand what one has been told, and as a result *do* what one has been told? Dennett says it is 'shrouded in fog' (consciousness being no longer explainable, apparently). He attempts to clear the fog by the manipulation of linguistic symbols. But in the midst of his explanation, the following occurs:

'The resource management techniques you are born with make no distinction between interior and exterior things.... The representation of features and things in the (internal or external) world become objects in their own right – things to be manipulated,

tracked, moved, hoarded, lined up, studied, turned upside down, and otherwise adjusted and exploited' (188-189).

Dennett puts the advantage of internalising 'the world' down to the flexibility of doing without the world *in thinking* (including imagining). But thinking has become self-justifying because Dennett wishes to account for the 'old' paradigm (i.e. thinking) without questioning whether it is a valid paradigm. Indeed this is so muddled, he states that the brain, which he terms 'resource management techniques', makes no distinction between interior and exterior things as if there needs to be interior *things* (unjustified reification) which are taken by the brain as the same as external things. He then moves to the classic supposition that words are labels which we learn to associate with things in the world.

'I'm suggesting that it's such initially stupid practices – the mere mouthing of labels [by a child]... – that could turn into the habit of representing one's own states and activities to oneself in a new way. As the child lays down more associations between the auditory and articulatory processes on the one hand, and patterns of concurrent activity on the other, this would create nodes of salience in memory.... These anchors of familiarity...could give a label an independent identity within the system' (198).

By 'the system' Dennett specifies, in none too explicit terms, the mind, although co-specified (equally inexplicitly as a causal function) is 'the child'.

These 'labels' and 'anchors of familiarity' 'finally reach...near-magical prowess...the *mere contemplation* of a representation is sufficient to call to mind all the appropriate lessons [of the learning process]. We become *understanders* of the objects we have created [i.e. words and sentences]' (200) (original emphasis).

This familiar account of the learning and activity of thinking takes place entirely within the 'fog', since it evades the distinction between brain and mind. The brain is physical stuff. The brain could have physical representations in it. But how could they be conscious?

Dennett claims his book is about consciousness, so he introduces the notion that representations, i.e. physical stuff, are *contemplated*. Contemplation is what a conscious human does.

But consciousness (and its contemplation) and physical representation are different. How could sentences be contemplated in the head except by there being physical sentences

written in the head ('things') which, by some 'magical' means (Dennett's own word), get *seen* and thence understood by the conscious subject contemplating them? This is not only impenetrable, it is Dennett's own Cartesian Theatre which he has deplored.<sup>112</sup>

Moreover, it seems to betray his notion of the 'I' as the *centre of narrative gravity* which, hypothetically, does *not* contemplate because it is not (i.e. really) contemplative consciousness. (This was his proposal in 1991, which appears left behind by 1996. We have commented on Dennett's vague use of terminology already (§4.2).)

By introducing the notion of representation, Dennett is trapped in the (intractably) dual nature of his explanation, between consciousness and physical stuff. Little wonder he makes the supposition that the brain (the organ of 'resource management techniques') does not distinguish between the internal and external – a distinction which, on the contrary, the brain cannot help but act upon to *be* the control system of the organism in the world.

Dennett does not explain either why there should be, or how there could be thinking. I.e. he does not explain why it should exist beyond causal electro-chemical processes. To say: 'Every agent has to know [or understand] which thing...it is' (1991; 427), knowing thus raised above electro-chemical causality, is tautological, the very kind he has criticised as we noted at §3.3.1. ('If your model of how pain is a product of brain activity still has a box in it labelled "pain," you haven't yet begun to explain what pain is.') But again, this is in 1991.<sup>113</sup>

Moreover, he does not explain what *communication* is – what physical processes take place when one individual talks to another, and how physical causality between organisms occurs as a result. In 1991 he says: 'These strings...of narrative issue forth as if from a single source...: their effect...is to encourage [an audience] to (try to) posit a unified agent whose words they are...*a center of narrative gravity*' (1991; 418). Apparently, the function of communication is also a kind of deception (but why?). By 1996 it has altered and become 'foggy'.

Thus, both accounts fail to *explain* language as thinking *or* communication.

---

<sup>112</sup> In a similar vein, Douglas Hofstadter's book, *I am a Strange Loop*, is wholly ontologically muddled. In Chapter 19 titled 'Consciousness = Thinking', he states: 'Consciousness is the dance of symbols inside the cranium. Or, to make it even more pithy, consciousness is *thinking*. As Descartes said, *Cogito ergo sum*' (2007; 276). He offers no account of what consciousness or thinking are in plausibly physical terms.

<sup>113</sup> In 1996 he says of the need for knowing: 'You couldn't sign a contract with somebody else's hand, mistakenly thinking it was your hand' (206). But why could you not sign it without knowing? I.e. by causal processes which are not conscious? Knowing cannot be justified this way.... His later notions of neural celebrity and fame defy intelligible comment.

### 7.2.2 Chomsky's approach to thought

There is more to be said about thought. Noam Chomsky is credited with revolutionising linguistic theory. In the third edition of *Language and Mind*, he quotes approvingly Ian Tattersall. Language is 'virtually synonymous with symbolic thought' (2006; 177). He continues quoting François Jacob: 'The role of language as a communication system between individuals would have come about only secondarily.' And further, now quoting Salvador Luria: 'Communicative needs would not have provided "any great selective pressure to produce a system such as language" with its crucial relation to "development of abstract or productive thinking".' He continues quoting Jacob again:

"The quality of language that makes it unique does not seem to be so much in its role in communicating directives for action" or other features of animal communication...but rather "its role in symbolizing, in evoking cognitive images," in "molding" our notion of reality and yielding our capacity for thought and planning, through its unique property of allowing "infinite combinations of symbols" and therefore "mental creation of possible worlds," ideas that trace back to the seventeenth-century cognitive revolution.'

Unfortunately, no explanation is given of what thought is, or how 'symbols' or 'cognitive images' exist, undoubtedly consciously. The account is thus ungrounded and tautological. 'Reality', apparently, is moulded by 'cognitive images'. But what is the relation of real reality to moulded reality, and who *has* the moulded reality – and *why* is it cognitive? What is the point of 'possible worlds' to be thought about, when we do not know what thought is? And why would 'communicative needs' *not* have produced 'a system such as language', when what language is has not been identified beyond the wholly obscure notion of thought?

Thought is simply a slogan, for no engineering account of how it works is provided. Chomsky may be right that the cognitive revolution goes back to the seventeenth-century. That is not an advantage: it is *exactly* what is wrong with it.

Chomsky is not fazed by thought being unexplained. He has a ready answer to those questioning whether mental states (e.g. thought) exist in the physical universe. 'The basic

contention of Priestley and other eighteenth-century figures seems uncontroversial: thought and language are properties of organised matter – in this case mostly the brain’ (2000; 115).

In other words, the problem does not enter his consideration. As far as he is concerned, ‘Lacking a concept of “matter” or “body” or “the physical,” we have no coherent way to formulate the issues related to the “mind-body problem”’ (110). (This is evident in his use, over his career, of the conceptually fudged expression ‘mind/brain’.) His justification is that Newton overturned the mechanical universe, i.e. the genuinely materialist view, by introducing occult forces, particularly gravity. He states:

‘In their correspondence, Leibniz and Huygens condemn Newton for abandoning sound “mechanical principles” and reverting to mystical “sympathies and antipathies,” “immaterial and inexplicable qualities.” Newton seems to have agreed. The context of his famous comment that “I frame no hypotheses” was an expression of concern over his inability to “assign the cause of this power” of gravity, which so departs from mechanical causes’ (108-109).

But Chomsky’s attempt to hold onto his thinking-version of language is not so saved. The key factor about physicality is not substantiality or materiality simplistically (or mechanically) conceived, but its behaviour under laws. The grasp science gets on the physical universe derives from the lawfulness it finds there – signified by its representations of it (categories-of-the-world). That was the significance of Newton’s discovery of gravity in relation to bodies (the units of the theory), whatever were the underlying causes of its effects. (Newton’s lacuna was moved forward by Einstein’s theory of *general relativity*.)<sup>114</sup> The problem with mental states is their failure (1) To be so appropriated, and (2) To relate to any feature of the physical universe (i.e. as lawfully expressible units) as so specified.

But even more basic than that, the question Chomsky does not address is: How can we have definitive knowledge that there are mental states, for we cannot claim knowledge of anything (cf. Chapter 4)? It may well be that we do not know what physicality is; but what we can say about it, i.e. endorsing that we *do* exist in a physical universe (as a reasonable assumption), is that the scientific approach has classified its elements as lawful. No scientific

---

<sup>114</sup> For example: ‘When you drop a glass, you can think of the event in terms of the earth’s gravitational field pulling on the glass, or, using Einstein’s more refined geometrical description, you can think of it in terms of the glass’s sliding along an indentation of the space-time fabric caused by the earth’s presence, or – if gravitons do indeed exist – you can also think of it in terms of gravitons firing back and forth between the earth and the glass, communicating a gravitational “message” that “tells” the glass to fall toward the earth’ (Greene: 2004; 255).

lawfulness for thought or feeling or sensation as necessarily (or predictably) *causal* exists. There is not even a specification for them as physical properties.

### 7.2.3 Thought *via* meaning and understanding

The mentalist problem with language as thought is how to describe the units – words, sentences – in terms of their causal properties. Language causal properties have two founding terms: *meaning* and *understanding*. Meaning relates to language itself and is the being of its causality; understanding relates to the subject for whom language exists and how the subject employs its understanding of meaning to be causal – even if this is no more than thinking in language.

We have run up against meaning and the problems with it (§§3.1.3 & 3.1.4). Meaning has been conceived in terms of sense and reference, after Frege.<sup>115</sup> A term, e.g. ‘water’, and our belief denoted by the term, involves *sense* (the causal psychological modality, or internalism), and *reference* (what the term or belief connotes, or externalism).

We ended our previous discussion at Slezak’s proposal that, to avoid the referential issue with meaning, the (mental) representation should not be supposed available to an observer and judgeable in terms of its truth – e.g. whether ‘water’ relates to H<sub>2</sub>O or XYZ – but should be considered functional entirely within the system in which it is operative. But this led us to the McGinn position that conscious cognition then becomes impenetrable to *our* conscious cognition. As a product of the brain our cognition cannot be interrogated by us as that very brain product. Not only is cognition as consciousness a mystery, consciousness itself is under suspicion (certainly in the Slezak version).

Since meaning depends upon consciousness (i.e. that which gives knowledge), what *does* ‘meaning’ mean? Evidently it means nothing. For meaning, as a technical term (as opposed to our expressive use of the word), is predicated on the notion of mind.

On the other hand, while substitution of one word by another (synonymy), or by a description of a word in other terms, is still possible – which for some *is* the meaning of

---

<sup>115</sup> ‘It is natural, now, to think of there being connected with a sign (name, combination of words, letter), besides that to which the sign refers, which may be called the reference of the sign, also what I would like to call the *sense* of the sign, wherein the mode of presentation [i.e. knowledge] is contained’ (1970; 57).

meaning<sup>116</sup> – the substituting or elaborating remain entirely within the rejected mind construct. In other words, there is nothing behind or within words except their existence in a physical sense, which itself is inaccessible as brain-sign, but nonetheless requires explanation (as we shall come to).

\*

However, a position which maintains meaning to be meaningful must rely upon the notion of understanding. For only *via* a conscious understander can meaning be meaningful. We find in the literature the extraordinary situation that meaning is debated while the understanding upon which it depends is denied explanation. Here is Paul Horwich.

‘Language is an expression of thought, a means by which statements are made, questions asked, instructions are given, and so on. Therefore to *understand* a language – to know the significance of its expressions – is to be able to tell which thoughts underlie their use: to know what is being asserted, asked, demanded, and so on. Thus the meaning of a sentence is the propositional character it expresses, and the meaning of a word is the element of such a character (i.e. the concept) that it expresses’ (2003; 137).

Horwich, in this passage, intends to reconcile this ‘natural’ conception of meaning with a Wittgensteinian ‘meaning as use’. He continues, therefore:

‘The concept DOG is most directly identified as the concept that is normally expressed in English by using the word “dog”.... On that basis, after the appropriate investigation into the nature of that shared use [language as a shared medium], we will be led to a characterization of DOG as that entity whose engagement by the speaker’s mind is manifested by his deployment of a word with use regularity.... Thus the natural view that the meaning of an expression is the concept it expresses is quite consistent with the further claim that such concepts are identified by the use regularities of the words that express them’ (137-138).

But this explanation by Horwich depends upon the fundamental elements he does not explain: thought, understanding, the ‘entity concept DOG’ and its method of engagement

---

<sup>116</sup> As with Quine. ‘Once the theory of meaning is sharply distinguished from the theory of reference, it is a short step to recognizing as the primary business of the theory of meaning simply the synonymy of linguistic forms and the analyticity of statements; meanings themselves, as obscure intermediary entities, may well be abandoned’ (1951; 22).

with ‘the speaker’s mind’ that allows him to deploy the word ‘dog’ so expressing the concept (cf., by contrast, §7.1.2).

In other words, although Horwich claims he can reduce the semantic features of language to non-semantic *use*, thus placing them in the physical world which he rightly regards as obligatory, his method is as ineffective as Pinker at §6.13 concerning the reduction of (heard) music. For he offers no explanation of what language *is*, what it does, and how it works which does not depend upon the assumption that language (and its elements) is how the natural disposition takes it to be, i.e. that we understand its meaning by our psychological capabilities.

On the contrary, to grasp how language works in the physical world, we must divorce it from all psychological concepts which, for Horwich, are still assumed. In his later book he states: ‘What ultimately constrains our theory are vocal sounds that a person produces, the circumstances in which they are produced, and the characteristic reactions to sounds that are heard. Such facts are plainly relevant and uncontroversially observable, insofar as they presuppose relatively few of the psychological hypotheses whose correctness might be at issue’ (2005; 176). Unfortunately there are *no* psychological hypotheses (including sound and hearing, and uncontroversial observation) that are uncontroversial, as we have demonstrated.<sup>117</sup>

This illustrates, as we have said previously, that there can be no solution to what the brain does and why by theories focussed on some parochial topic without addressing the overall construct. (That construct entails the notion of consciousness essentially, from which *all* psychological properties derive.) Yet this appears to be the general state of the art.

### **7.3 The social factor & learning language**

We now move towards our account by considering Philip Lieberman’s instructive book, *Toward an Evolutionary Biology of Language* (2006). He makes these two points.

---

<sup>117</sup> In his most recent book Horwich intends to ‘present a broad and unified “deflationism” that encompasses language, thought, facts, knowledge, and the relation between them’ (2010; v). He concludes from this that we can ‘find that every one of the theoretical positions typically adopted in [long standing debates in metaphysics and epistemology] is misconceived.’ His method is to provide a “no-answer” answer to the perennial question, “What is truth?”, and a meaning as use approach in which ‘the sense of each word-type is given by its basic patterns of deployment rather than by its association with some feature of the non-linguistic world.’ While his intention, to demystify language, is laudable, his method is too local to be genuinely effective.

1) 'The body of evidence...suggests that children acquire words and syntax by means of associative learning, imitation, and the subtle social cues that parents and caretakers use to convey displeasure to children or to correct errors. Human beings are intensely social animals and normal children are sensitive to these cues at an early age' (239).

2) 'Social factors have been cited as driving natural selection that increased neocortical size, thereby enhancing lexical and syntactic ability. It is hard to disagree with the general premise' (*ibid.*)

Could there be language, or indeed human beings, unless they were collective animals? Is it remotely possible that organisms of the adaptive capacity of humans could have developed solipsistically, or that they could have evolved the capacity for thinking (in their solipsism) with language? These are rhetorical questions.

The natural disposition supposes that individual human beings communicate with each other by language which carries information that can be used in thinking.

The scientific disposition does not make this supposition. Although it accepts there are separate organisms, they also function by processes of collectivity which use the physical intermediaries of compression waves and electromagnetic radiation. They can alter each other's behaviour to mutual benefit as the super-organism.

Thus the *capacity* for 'learning' language is not added to by surrounding instructors, like parents and caretakers. Language is built *de facto* within the super-organism as part of its behavioural structure as the unit of the communicated.

This is not an argument against Lieberman. Indeed, he explicitly differentiates his position from Chomsky, his mentors and legacy. Our position extends Lieberman's by escaping the notion of consciousness, the significance of which Lieberman has not recognised.

\*

Is Dennett correct in his account of how we learn words? He says (§7.2.1):

'I'm suggesting that it's such initially stupid practices – the mere mouthing of labels [by a child]... – that could turn into the habit of representing one's own states and activities to oneself in a new way. As the child lays down more associations between

the auditory and articulatory processes on the one hand, and patterns of concurrent activity on the other, this would create nodes of salience in memory.... These anchors of familiarity...could give a label an independent identity within the system.'

This cannot be explanatory because Dennett confuses different processes and elucidates none. Undoubtedly those who use language teach those who as yet cannot. But the physical processes involved cannot be simply described by employing the word 'child' or the 'mouthing of labels' or 'representing to oneself', for this implicitly assumes psychological concepts which have not been justified.

Adult brains train child brains. The process of pointing to an object and saying the word, e.g. 'water', may be involved. But as commonly agreed, the word 'water' is arbitrary. In what way does *it* exist? More accurately, what is being trained?

The adult brain trains the child's brain in its causal states, not its knowledge or use of words. (There is no mind housing a repertoire of words.) In so doing, shared categories-of-the-world are established between them, e.g. water as brain-sign in its various modalities, 'seen', 'felt', stamped on in puddles. (Of course much of this training is carried out independently by the child.)

In terms of language, these categories-of-the-world become designated by other physical processes. The physical structures of compression waves, transduced from the adult's to the child's brain, constitute the *physical* process of teaching and learning, relating words to categories-of-the-world (§7.1.2). So the child's brain-sign 'hearing' of the word 'water' can coincide with the brain-sign modalities of the causal orientation towards the presence of water (though not caused by those modalities) as learned. The training of these networks which emerge as the brain-signs of language is of course complex. They associate with other networks, and develop over time. Thus subsequent 'hearing' of the word 'sea' can be associated with many different 'images' (etc.) of the sea, even in the one brain, from the various networks. What does not happen is 'representing one's own states and activities to oneself'. There is no self to whom one's states and activities could be represented. Nor are there 'nodes of salience in memory', for there is no memory of an 'independent label identity' – nor indeed the sea itself. The brain's orientation is causal, which results in the relevant brain-sign occurrences of both categories-of-the-world and words.

#### 7.4 Lieberman's attack on Chomskian language

As mentioned, psychology regards the human psyche as composed of faculties of mental type, e.g. seeing, hearing, feeling...and language. The physio-ontological problem, as we have seen, lies in the words of John Heil.

‘Even if we knew that neural tissue arranged in a particular way yielded a feeling of pain, the reason this arrangement yields pain rather than some other feeling (or no feeling at all) remains an utter mystery.’

We will address this again now for language. We begin with the critique on Chomskian language undertaken by Philip Lieberman.

Lieberman's central contention is there no language faculty (or module) performing the sole linguistic role, distinguishable from other activities of the organism. He refers to the converse as the ‘standard’ view, found with such theorists as Chomsky, Fodor and Pinker. Further, Lieberman proposes that language is an action and replicatory of, or having an equivalency to, other actions.

Specifically, he rejects Chomsky's notion of a Universal Grammar. What has become for Chomsky a crucial feature of language – that all languages are determined by an innate set of rules or syntax which are essentially genetically encoded and allow the permutation of words into an infinite number of sentences – is a misguided differentiation of language from ‘motor control and other aspects of animal and human behavior’ (2006; 5). As he says, ‘The chance events and evolutionary changes that yielded human linguistic and cognitive ability are apparent in the human brain and body’ (6).

This implies both that language does not suddenly spring into being as a unique faculty, and that it is not the unique differentiator of intelligence between animals and us. This is reflected in the words of Euan Macphail from experimental studies.

‘Supposedly unintelligent animals [who do not have language] have turned out to be intelligent.... Association formation [is] equally efficient in all vertebrates [and] associations clearly play a central role in many (if not all) complex tasks’ (1998; 120).

A particular aspect of the parallel between language and action is quoted by Lieberman from Karl Lashley, who he takes as a progenitor of his views.

‘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’ (2006; 12) (Lashley: 1951; 113).

We have already pointed to this (§§5.3.2.7 & 6.8).

Another aspect of Lieberman’s critique is the neural one. He rejects the notion that language is focussed in the ‘classically’ supposed areas of Broca and Wernicke. This is because, not least, studies have shown that subcortical areas involved in motor control and cognition are also activated. As he says:

‘Evidence from hundreds of independent studies that span three decades show that different regions of the neocortex and different subcortical structures are specialised to process particular stimuli, visual or auditory, while other regions perform specific operations that regulate motor control.... However, these local operations form a set of neural computations that link together in complex neural circuits, yielding actions such as walking, pushing a button, or comprehending this sentence’ (2006; 158-159).

A region of particular significance is the basal ganglia complex, which ‘includes the putamen, caudate nucleus, global pallidus, with close connections to the substantia nigra, thalamus, and other structures. This system essentially acts as a sequencing engine’ (169). Damage in this area, therefore, inhibits movement *and* directed cognition *and* language use.

\*

However, Lieberman demonstrates a confusion in specifying the link between ‘walking, pushing a button [and] comprehending this sentence’. For walking and pushing a button are physical actions, whereas comprehending a sentence is something else. Cognition, which we may take as an intelligent response to the world in terms of sequenced physical activity, does not require *comprehension* of language.

On the other hand, as Macphail notes, language as communication is of primary significance for human cognitive superiority over other animals – including increased cortical size. Indeed, Lieberman makes much of the power of language in the *physiology* of speech.

‘The process that allows human speech to attain [a remarkably high] rate of transmission involves encoding phonetic events into a speech signal that consists of units whose minimum length is a syllable. Human listeners appear to decode the speech signal by using a neural system that inherently takes account of the constraints of speech production. The listener’s perceptual system “knows” the set of encoded signals that reflect speakers’ execution of motor acts that correspond to a series of individual phonetic segments’ (2006: 98). ‘The structures of the human brain that are involved in producing speech are activated when we listen to words’ (100).

What does this tell us? Not that we hear language comprehendingly, but that the causal inter-communication of human brains (as with other so-called mirroring activity, cf. §6.5) is a purely neural function in the domain of action. And this has nothing to do with the brain ‘knowing’ anything, which Lieberman hints at by his use of quote marks.

But in addition we can infer that the ‘hearing’ of language is the signification of the causal status of brains *from* that neural inter-communication, wholly in line with brain-sign theory generally (§7.1.1).

\*

Yet although Lieberman designates language as an action, and finds its source in diverse structures of the brain, and even puts ‘knows’ in quote marks concerning the brain’s activity, he retains the notion that the brain has a dictionary of words.

‘When I hear, read, or think of the word “elephant,” I can relate it to an image in my mind’s eye of an elephant or any number of properties of an elephant.... Martin and his colleagues found that the primary visual cortical areas associated with the perception of shape and color are activated when we think of or read the name of an animal such as an elephant’ (34).<sup>118</sup>

As already stated in this chapter, and is true generally of introspection, there is no way to establish that the ‘heard’ word *causes* some other aspect of brain-sign. Lieberman, and the original investigators, misinterpret the data because their theorising is not founded upon a plausible brain science.

---

<sup>118</sup> Reference is to: Martin *et al.* 1995a & 1995b.

What causes both the ‘hearing’ *and* the ‘image’ are structured compression waves or electromagnetic radiation transduced into the brain, and interpreted by the brain as ‘hearing’ and ‘reading’, thence associating with other networks to produce concurrent ‘images’, i.e. categories-of-the-world.

What has been supposed as thinking, when not caused by external factors, is simply neural self-activation, as repetition or reformulation, with resulting brain-sign occurrence (§6.10). Co-instantiated with the brain-sign word ‘elephant’ is the ‘image’ of an elephant. Thinking of the word ‘elephant’ does not cause the image of an elephant. There is no thinking *qua* mental state.

## 7.5 Language as action

We can now state what language is. Language is a bodily action caused by the brain, by which one brain can penetrate into the brain of another directly and remotely, in the unit of the communicated, so altering its causal orientation. As brain-sign, language (words etc.) are then the communicative condition of that brain’s causal status.

Resulting action may occur, it may not occur immediately and it may not occur at all: but the means exists for the possibility of resulting action.

Contrast this with the lion hunt. The movement of lion A can cause a move by lion B. But the entire premise of brain-sign depends upon the possible ambiguity of lion A’s move, which has generated the occurrence of the internalised sign in lion B (brain-sign) of lion A’s move as explanation of its (lion B’s) move. With language, ambiguity is reduced. For what is transmitted is direct command, explanation, request, etc. But not as heard words or sentences. The transmitted command, etc., occur as structured compression waves (or electromagnetic radiation) transduced into the brain. The brains of requisite organisms are organised to facilitate this kind of *causal* effect. ‘Hearing’ a command is brain-sign ‘explaining’ for others the causal state of the brain resulting from the transmitted compression waves. (Peirce gets half of this almost correct, §5.2.1.)

How can there be different languages meaning the same thing by different words (without such inexplicable mentalist foundational notions as *Mentalese* or *state-space*

*semantics* underpinning them)?<sup>119</sup> In fact the arbitrariness of words is irrelevant because words as ‘heard’ have no causal function. As brain-sign, however, they signify within an established framework of communication which, itself, is invisible to (us as) brain-sign. (Thus we have a more fundamental neuroscientific explanation of what, for Horwich after Wittgenstein, is ‘meaning as use’ in the patterns of word-type deployment, §7.2.3f.) As Saussure correctly says, but without appreciating its full import: ‘Of all the social institutions, language is the least amenable to initiative. It blends with the life of society, and the latter, inert by nature, is a prime conservative force’ (1970; 49).<sup>120</sup>

### 7.5.1 How language functions

Words appear to be objects ‘out there’. In brain-sign terms they might seem to be categories-of-the-world because, for example, there they are on the page like a table or chair in the room. But this completely fails to grasp the function of words (and language) as a brain-sign category. Of course the marks on the page are a category-of-the-world. But before we ‘see’ or ‘hear’ *words*, our brains are already causally orientated toward the world in a way completely invisible to us (as the brain-sign of the words), and this accounts for the fact that we *appear* to know what a word means, and that we *appear* to understand how to use it. The sense of ‘meaning’ and ‘understanding’ are the explanation (as brain-sign) of the causal orientation we already are which the impact of the structured compression waves or electromagnetic radiation have caused in us from the ‘word’ source. That *sense* of meaning and understanding have no depth: they are a transience of brain-sign itself (cf. §7.1.2). (And are categories-of-interaction, though with language, not the world. They are, however, displaced, cf. §6.12.3.)

Again Saussure has half-grasped this when he says:

‘No individual, even if he willed it, could modify in any way at all the choice [of the sign] that has been made; and what is more, the community itself cannot control so much as a single word; it is bound to the existing language’ (1970; 47).

---

<sup>119</sup> For *Mentalese*, see e.g. Fodor (1975), and subsequent publications; for *state-space semantics*, see e.g. Churchland (2007).

<sup>120</sup> Why is there indeterminacy in translation? (Quine: 1960). Why is it impossible to state with absolute precision what a word in an unknown language means in our language? Because what is causal in language is not codified in either the sense or reference of words. The search for the psychological or referential status of words *as causal* is therefore futile.

The lack of freedom, however, exists in the ontological structure of the operation of language. This is further (but unclearly) endorsed by Saussure when he says that ‘language always appears as a heritage of a preceding period’ (*ibid.*). The *appearance* of ‘a heritage of a preceding period’ is really the physicality as which language functions *as us*, and to which *as such* we have no access. Of course it derives from past ‘practice’; but the inability for change resides in the physical condition as which individuals and society exist – which is contrary to the way society supposes it exists in the natural disposition. Saussure’s statement that no act of will could modify the choice of sign acknowledges (without grasping it as such) the impotence of the notionally mental in the condition of language.

### **7.5.2 Communication with animals**

Since it is incorrect to suppose *we* understand words, neither do animals. But Lieberman says: ‘The circus dog studied by Warden and Warner understood 50 English words. A contemporary border collie raised in German-speaking homeland understands more than 200 words’ (2006; 36).<sup>121</sup>

Transduced compression waves from humans may cause dogs to act in certain ways. The associated brain-sign occurrence in dogs and humans may be the degree of neural communication. But dogs no more have a mental life, to which they have knowing access, than do we. Thus Lieberman’s use of the word ‘understand’ is colloquial, not scientific.

However, the communal interaction of humans with dogs is greater than with cats because many dogs can interact verbally (i.e. causally), albeit essentially asymmetrically. The lack of reciprocity (which is not absolute; we too can react to canine ‘sounds’) results from the absence in dogs of adequate vocal mechanisms neurally driven.

### **7.5.3 Intelligence lies in action**

The complexity of human language is not simply attributable to what can be communicated. Linguistic complexity results from the actions of which humans are capable. It is incorrect,

---

<sup>121</sup> Referred studies are: Warden & Warner (1928) and Kaminski *et al.* (2004).

therefore, to propose that human intelligence *results* from language. Language, and what humans can do, must co-extend to a degree, for language *is* an action.

In other words, we are not more intelligent than apes because we have language. Intelligence lies in what we can do, and we can do much more than apes (but not comprehensively). Communication by language supports intelligence in action, particularly collective action, because it *is* an action. Without the notion of language as action, its existence and function is incomprehensible. Bernard Comrie and Tania Kuteva state that:

‘Our approach predicts that human language must originally have been much simpler, and highly functional, but that over the millennia since that time, various kinds of elaboration have developed. More precisely...we assume that certain complexities of all or many currently attested languages were not present in earlier human language’ (2005; 202).

Highly likely. But language cannot have developed as an isolated faculty in complexity. What job would it be doing? Linguistic complexity arises because the communication of communally acting organisms extends the power and complexity of action. It is the failure of Comrie and Kuteva to appreciate this that lands them here:

‘When speaking about the way language evolved in human prehistory, it is highly implausible to assume that humans developed language because they *needed* it in order to adapt better, and faster, to their environment’ (185).

This will not do because the authors offer no account of what language is and does *via* which its adaptational impact could be assessed.

## **7.6 The operation of language**

As the direct control mechanism of *other* brains, language has evolved. There are two aspects to this: (1) The neural production of language; and (2) The neural reception of language.

The neural production of language (1) functions in different modalities. In the natural disposition, explaining my condition to another is the attempt to get them to understand it. If I am anxious about some future event, I will be informing them of my anxiety and why. In the scientific disposition I am not a person so informing another. My brain, *via* language

(structured compression waves), alters another's brain so that, in the unit of the communicated (2), the other will act taking my condition into account. I 'hear' my words, and so does the other. But the brain-sign of 'hearing' is 'merely' the mutual neural communication.

The bio-functional cause of my speaking is to alter another's neural status. As brain-sign I have no access to that cause. I may 'feel' an overwhelming need to communicate my anxiety (as category-of-interaction). But my sense of that need is another aspect of brain-sign, explaining why I do speak. Indeed, I might say 'I just have to tell you....' , thus more completely conditioning the other's brain.

There are a very large number of ways in which communication can take place because there are numerous ways humans can act. Another modality is command. I may tell you to put on a jacket because it is cold outside. What causes you to put on the jacket (if you do) is the modification of your neural status by the compression waves issuing from my mouth. 'Hearing' my command has nothing to do with whether you do or do not put on the jacket.

My command, as the natural disposition has it, could be suggestive, advisory, determinant, etc. This demonstrates, even within one modality, the subtlety of human causal interaction *signified* by these descriptive words. Each reduces to the biology of our neural reciprocation and is attributable to our biological status which, in command mode, may be signified as close friends, acquaintances, a warden and his prisoner, and so on.

### **7.6.1 Neural collectivity**

From the stance of the super-organism, it can be no surprise that our heads are filled with the causal impact of others. The supposition that we have an identity hermetically distinguishable from others is absurd. Our causal neural structure is an amassment of a history woven from a collectivity unavailable to our consideration, for we are constructed in and from that collectivity. The 'self', recall, is the transient brain-sign representation of the body in its causal immediacy (§6.9).

Because the natural disposition promotes the notion of the individual, our communal being lacks emphasis: methodologically it is taken for granted. But how else could we communicate except as a *physical* collective?

As Brutus and Antony speak in the Forum after the murder of Caesar, they illustrate quintessentially what might be termed public relations, in which the crowd is drawn from one position to another.

As Brutus has it:

‘If then that a friend demand why Brutus rose against Caesar, this is my answer: Not that I lov’d Caesar less, but that I lov’d Rome more.’

As Antony has it:

‘But here’s a parchment with the seal of Caesar;  
I found it in his closet – ’tis his will.  
Let but the commons hear this testament,  
Which, pardon me, I do not intend to read,  
And they would go and kiss dead Caesar’s wounds  
And dip their napkins in his sacred blood.’

As we read these speeches (with their complex rhetoric), or hear their entirety in the theatre, the biological *how* by which we are taken over by their impact (as are the crowd) is invisible to us. Shakespeare had no conception of the physical mechanisms by which speech operates. What he portrayed so effectively, however, was the mutability of human *behaviour* (the baying crowd) by the physical processes of language.

## **7.7 The nature of language, and our fundamental limits as biology**

We have drawn attention to two significant aspects of language. The first is its parallel with other states of brain-sign (§7.4); the second is that (as Saussure hints) we do not know what language is (§7.1.2). These require more elaboration.

If we suppose perception is input to the cognitive process we have the problem of specifying what perception is and how it *can* so function: i.e. what is its effective ontological nature? If the brain phenomenon is brain-sign there is no such problem.

As brain-sign, the 'image' in the lion's brain in structure and content derives from the brain's causal orientation (§5.3.2.7). Hypothesising that (in some manner) the image 'contains' the gazelle, the other lion and the terrain, it can be considered both semantically and syntactically as if it were a language. The brain (in normal operation) does not mistake the other lion for the gazelle, nor the scrub-land for either. In other words, what these features signify (as semantics) are well-defined in terms of the lion's causal orientation to the world. The structure of the brain-sign content, i.e. the relative positions and movements of its elements, we may regard as its syntax. Semantics and syntax therefore signify the *why* of the lion's causal activity.

But although we can describe the lion's brain-sign condition in these terms, we have no way of grasping it beyond stating it. We have reached the limit of the explanatory process. For brain-sign *function* in content and structure is a fundamental property of the universe. This must evidently be the case, for it is the being of the existence we are.

So it is with language, for language as brain-sign is also the content and structure of the neural interpretation of the brain's causal orientation. But it functions in a different way from the image in its semantics and syntax. The causal function of what results in 'heard' or 'seen' language is designed to alter another's brain. These two ('image' and language) may be termed *implicit* and *explicit* communication.

A fundamental difference between explicit (language) and implicit ('image') communication is that the syntax of language is (once acquired) rigid by comparison. It functions in what we may describe as static rather than dynamic communication, as with the bee dance. Indeed it must, so to be effective. Whether or not language works in a given case, its results have evolved to be precise.

In what way have we reached the limit of the explanatory process with language?

When McGinn states that: 'We need to cultivate a vision of reality (a metaphysics) that makes it truly independent of our given cognitive powers as a proper part.... A deep fact about our own nature as...embodied consciousness is thus necessarily hidden from us'

(§3.1.4), he makes two crucial errors. He did not appreciate that his statement, though insightful, was both a self-contradiction and an obscurity.

Consciousness, being knowledge, is incapable of not knowing itself by definition (self-contradiction); or better perhaps, a ‘not knowing’ indicates we do not know what consciousness is (obscurity), i.e. we cannot suppose we *are* conscious if we cannot explain what consciousness is (*via* itself). Nothing is explained by saying consciousness is ‘embodied’. What limits our cognition is the nature of its existence (ontology) – which is certainly embodied (the neural brain), but not conscious. The brain reacts to the world by its physical neural conditions. The brain knows nothing of the world; but it can signify the cause of its causal status – brain-sign.

That we appear to have knowledge is brain-sign operating successfully. The ‘deep fact about our own nature’, however, is that *everything* is, as it were, ‘hidden from us’ because what we are is brain-sign, not consciousness.

Just as our brain-sign ‘images’ are one explanation of our brain’s causal status, language is another. When I say to my friend, concerning the lime tree (§6.12.1): ‘Of course, you’re right; I don’t know what’s come over me today,’ I am not imparting knowledge of *my* physical condition. My brain is signifying the cause of *its* condition. What brain-sign theory reveals is that I am declaring my *ignorance* about my condition, both before and after. For although I may now sense that my friend’s words have altered my condition, I have no idea (I do not know), from the condition itself, what was so before and is now subsequently.

There is nothing more that can be discovered about this means of communication because its being is a fundamental property of the universe and cannot be self-revelatory. There is nothing or nobody to which or to whom it could reveal itself. As Slezak rightly says (§3.1.5):

‘Apparently talking about the meanings of inner representations, Block asks “what is it to grasp meanings?” and “what is it for the brain to grasp meanings?” The idea of a grasping or understanding here is exactly the wrong one in relation to internal representations which are precisely not grasped or understood.’

We may or may not locate where in the brain language is ‘heard’ (the virtual array). But this would not explain ‘hearing’ to us, for our ‘finding’ the location would be the way our brain is causally orientated to that location; our shared brain-sign *of* the location would not be

knowledge *as which* ‘hearing’ is understood or explained. Indeed, even the finding of the location is not knowledge.

Thus when I think I know something because the words or formulae occur in my head, e.g.  $E=mc^2$ , I mistake my condition. There is no thinking as some kind of cognitive (i.e. knowledge) state for me as mental subject. Pushing further than Slezak – because we have a coherent model – we see that the way ‘we’ operate in the universe is other than the supposed self-revelation of causal consciousness.

In other words, there are, for us, two levels of remove from the notion of knowledge as the divine revelation of the universe. The first is that, as brain-sign, we are already an interpretation by the brain of its causal conditions – which is not knowledge. The second is that brain-sign, which we are, cannot identify the being of its own state, for it cannot self-reveal its own condition.

## 7.8 A recreational coda

Georges Rey says (§3.2.2): ‘My favourite examples are jokes: how could one hope to explain regularities in people laughing at jokes without adverting sometimes to their content?’

What Rey means is that only by taking into account what the joke is about is it possible to explain its effect. (Although it is curious that he uses the word ‘sometimes’ rather than ‘necessarily’.)

Here is a classic Music Hall joke.

Man A: I say. I say. My dog’s got no nose.

Man B: No nose? How does he smell?

Man A: Terrible!

What Rey intends is that we could not appreciate this joke as *funny* (the mental state associated with the effect of a joke) without also appreciating its content: what a dog/nose is; what not having a nose means; what smelling is – *as* intentional mental states.

The classic formula for a joke (and we will consider only this form) is the turn of the unexpected. Freud wrote a book on jokes, and here is his explanation of them.

‘A preconscious train of thought is abandoned for a moment to be worked over in the unconscious, and from this it emerges as a joke. Under the influence of the unconscious it is subjected to the effects of the mechanisms that hold sway – condensation and displacement’ (1963-1964; 235).

In the case of the Music Hall joke we are first directed by Man A towards the lamentable state of the poor dog that is deficit in olfactory capability. Man B expresses his concern for the dog and its inability to smell. We are *then* expecting an answer to B’s question that relates to how the dog deals with its situation smelling-wise. But we are immediately jolted by Man A’s answer into a different context by the new information that the poor dog stinks. This makes us laugh, not merely because we are forced into an entirely and unexpected train of thought, but because our former sympathy for the poor dog (along with Man B) is brought up short by the fact that the dog is in fact appallingly malodiferous as expressed by the same Man A who initially elicited our sympathy for his dog.

Freud’s aim was *precisely* to avoid the problem Rey indicates. He wanted to explain how jokes worked without recourse to the singular notion of consciousness and its knowledge capacity. How *could* one react to a joke, as the turn of the unexpected, within the wholly transparent competence of knowing? And, moreover, when this involved a specifically *motor* reaction of a unique kind (laughter)? His method was to bury the mechanisms in the preconscious/unconscious which, by assumption, are closer to motor function and operate outside the knowing capacity. In other words, the forces which determine our behaviour in this context, even intentional ones, arise from a part of us which is not available to us consciously.

But Freud’s attempt was half-baked. The unconscious, as an explanatory domain, is useless because there is no explanation of the mechanism.<sup>122</sup> The question is: What actually happens with jokes as a biological operation in the physical universe? When Man A makes his initial statement (as structured compression waves), neural networks within us react in a causal manner. There are thus two states within us. The first is the causality which is highly complex, involving the reaction to the conditions of the joke being told, but as compression waves. Nowhere in us is there a mental replication of the physical world (as mental dog or nose, etc.) having a causal impact. But secondly, our causal orientation is interpreted by the

---

<sup>122</sup> This resulted from a failure of ontological scrupulousness, evident in Freud’s early *Project* (1950). He attempted to describe the functioning of mental states as if they were physical networks. But there are no mental states, or perhaps one should say, Freud did not make the ontological distinction which is crucial.

brain, resulting in the brain-sign state of the joke being told, of dog and noselessness; and this acts as the communicative medium with the joke-teller of our brains in the physical universe.

Since the initial line of the joke sets up our causal orientation, which also associates noses with ‘smelling’ and a ‘sympathetic’ *move towards* the poor dog (as our brain-sign state may already indicate), our causal reaction to Man B’s question is an already determined state of our brains (e.g. his is the kind of question we would have asked). Nowhere at this point in our causal reaction is there any *retraction* from a physical situation because of the physical impact upon *us* of malodorous-causing molecules in our olfactory organ, which our body would take as *to be withdrawn from*.

But Man A’s following line immediately triggers our networks such that there *is* a causal response, activated but not active, which *is* a retraction; and it is the juxtaposition of that *retractive* causal network with the previous *move towards* network which causes a motor response (laughter) and the *sense* of funniness (as the natural disposition has it) as brain-sign.

Both Man B’s question and Man A’s following line occur as us as brain-signs, with the reaction sharable by all concerned (‘funniness’ as a neural sign). But now all is explainable within the physical universe, where the notions of ‘sympathy’, ‘repulsion’ and ‘funniness’ translate into brain-sign states signifying the causal physicality of the juxtaposition of our neural conditions (*move towards* and *retraction from*), together with the motor response of ‘laughter’.

Funniness is not a kind of mental *existent* associated with the content of jokes which we have knowledge of, as Rey assumes. (Rey’s theorising starts from the wrong model, which is why it never escapes that model.) By eliminating the mental properties language is supposed to instigate (as e.g. heard content), brain-sign theory explains in the physical universe: how jokes work (in this representative example); what laughter is as a motor function (a neural conflict, as it were); and where the *sense* of funniness resides, i.e. as a category-of-interaction with (in this case) the physical structures of language.

Of course, the expressions ‘move towards’ and ‘retraction from’ require neurophysiological explanations which are, as yet, beyond current science. While in the case of *this* joke, specific categories-of-the-world and categories-of-interaction are invoked by the structured compression waves of the joke telling, their organisation as ‘move towards’ and ‘retraction from’ are systemically conflicting, a conflict which is operational in *all* such

jokes. It is that conflict that causes the physical laughter, which is not generated by the ‘content’ of the joke at all. Nonetheless, the arrival at the conflict, in the different networks of different jokes, disguises (as it were) their uniformity.

### **7.8.1 A review of the two forms of joke explanation**

What has language itself achieved in setting up this here presented alternative scenario? For we now have two explanations for jokes: what we might call the conventional (or mentalist) one, and the brain-sign version.

Language (by the ‘reading’ of this text) has caused a new theory to inhabit the brains of readers where before it did not exist. But according to brain-sign theory, neither theory can be designated as true. For they are both held as purely causal states of physical brains, which signify themselves at times as brain-sign.

Which is better? Of which can we be certain? In which can we believe?

Once brain-sign theory is grasped (or rather becomes our operative neural condition), the certainty of truth, and even the conviction of belief, are seen as inappropriate approaches to how we exist. The brain has no relation to truth, and the brain believes nothing.

In other words, the justification of one’s states by one’s beliefs could only occur if there were such a thing as belief. But that requires mental states to exist. According to brain-sign theory, mental states do not exist. Therefore one cannot justify one’s states by beliefs – because there are no mental states. Believing in beliefs is, therefore, a form of delusion (brain-sign theory would suggest). A claim to Godhood, or the megalomania to which humans are intrinsically destined (to which Nietzsche and Sartre have drawn our attention). The apogee of belief-supposition is knowledge.

Of course ‘we’ can *say* we believe things. But ‘saying’ is a communicative mechanism by the physical organ, the brain, the causality of which does not involve belief (cf. §7.1.1).

Similarly one *can* approve of brain-sign theory because it works better than mental state theories by expressing itself explanatorily as part of the physical universe. Mental state theory cannot.

However, whether one adopts brain-sign theory has nothing to do with whether one approves of it or not. 'Approval' is simply a collective expression of how we have acted according to our causal neural states to which approval has no access. Hence also consciousness.

## **The Theory of Brain-Sign**

### **8.0 The organism without consciousness.**

#### **Science and other human disciplines.**

Human disciplines (of ‘thought’) are a human creation. But we have redefined how a human being works. Therefore the interpretation of human disciplines will change. They do not result from a being who knows the world, as the theory of consciousness would indicate.

First we explore more of the conditions of brain-sign; then we consider the new interpretation. As indicated at the beginning of Chapter 6, brain-sign theory opens up vast new territories for investigation, and we can only skim the topic. The discipline chosen here is science.

### **8.1 Theories (leading to a consideration of science)**

The Shorter OED (1973) begins its definition of the word *theory* as follows:

1. ‘Mental view, contemplation’ from the Greek *theoria*, and Latin *theoria*: ‘contemplation, speculation, sight’.

It continues: ‘2. A conception or mental scheme of something to be done, or of the method of doing it.... 3. A scheme or system of ideas held as an explanation or account of a group of facts or phenomena.’ Etc.

It is evident the definitions depend fundamentally upon mental life, including language as contemplation, speculation as imagination, and sight. Moreover a theory is deemed to provide an explanation of phenomena somehow available, and facts as supposedly established or establishable.

One can hardly fault the OED for its definitions since it is operating within the culture to which it is beholden. But it may well be apparent now that human culture is wayward in its assessment of the physical universe, and the place of the human in it.

In the natural disposition we suppose we do have theories. They may be learned in science,  $E=mc^2$  for example, or they may be explanatory assessments of something that has

happened to us. ‘I forgot to pick up my keys because I was preoccupied with our missing cat.’ Indeed, if we had no theories (it seems), we would be able to do very little, for our capacity for action depends upon them.

But the *being* of the explanans (that which is the explanation or theory) in these accounts is itself not explained. The OED has no concern for that; neither, in normal life, have we. The ontology of our existence *as* theory, or the consciousness presupposed by us for theories, is not questioned. Science, however, puts them in question. How could it be that theories exist as conscious phenomena when consciousness does not exist? Answer: they cannot.

But brain-sign theory does not say that theories do not exist. On the contrary, they do so as the physical being of the human, not as the knowledge humans have *via* consciousness. How, then, do they so exist? (We will now refine statements in §4.3.)

The first mode of a brain theory (one might say), and upon which the brain acts causally, is its physical condition. But this bears no relation to consciousness, or theories as discursive statements. And it seems unlikely that the mode in which the theory exists causally *as* the brain could be discovered (to any great extent). This fits with the condition of a neural network in its connectionist response to an object in the universe (§4.3). For the theory is dispersed in the fabric (connections/synapses, and other elements of physicality) of the brain.

The second mode is brain-sign itself. Sight (for example) does not exist, but the brain-sign misunderstood as sight *is* a theory. It is a derivative theory. It is the brain’s ‘graphical’ *explanation* of its causal orientation in the unit of the communicated.

Thus we can readily say that lions have theories (§5.3.2.7). But they are not discursive theories upon which lions knowingly operate in their mental states (as the OED might propose). They exist as states of lion brains allowing action to take place (first mode), and communicative reciprocity for the super-organism (second mode).

A similar account can be proposed for language, and here we *are* talking about discursive theories. But the difference from mentalist notions is still evident. For the discursive theoretical sentence (for example), or the concepts it notionally invokes, are not causes of our actions. The causal scientific theory is embedded in the physical structure of the brain. It is taught to others by structured compression waves or electromagnetic radiation, which alter the other brain, orientating the organism toward the world in a causal way. It is

the derivative brain-sign state that is apparently discursive. (Although in a manner invisible to ‘science’ for, not least, nothing is visible, or as the OED has it, of ‘sight’, cf. §7.8.1.)

Thus what is indicated by the word *theory*, in brain-sign terms, is quite different from OED terms. Since brain-sign theory demonstrates that words do not mean anything (§§7.2.3 & 7.5.1), theories are physical states, and their nature and function depends upon how they are *as* physical states.

The only way a brain can respond causally to the world is by altering its physical states. Therefore the success or otherwise of its responses to the world depend upon the efficacy of its causal properties in relation to the world. The natural disposition might replace ‘success’ or ‘efficacy’ with ‘corresponding’ or ‘true’. But a brain cannot *know* what corresponds or is true. It is circumscribed by its physical nature, and its effects can never be tested as corresponding or true, for the universe *per se* is invisible to, or unreachable by the brain – and therefore ‘us’.

However, the success of science demonstrates that the brains of scientists work effectively in their physical properties. That does not result from a primordial reaction of the individual scientist’s brain to the universe. Extensive cultural pre-programming is necessary before any creative move forward is made, scientifically. No one knows what is involved in that cultural pre-programming, or what renders a particular individual’s brain ready and capable of making scientific progress, for brain science is in its early stages. Nonetheless it does happen, which demonstrates the effective embedding of our existence in action (in science not least) as determined by what we do *not* know, and can never *know* (cf. Pragmatism and Heideggerian analysis, §§5.2.1, 5.2.2). This addresses the point made at the beginning in Chapter 1: How is it that routine work on patients’ brains is carried out in hospitals across the planet when the status of the constructs upon which those operations take place (brain and mind) are not addressed, and their nature remains a Holy Grail?

Routine work can be carried out because the brain’s work involves no knowledge which depends on minds, for minds do not exist.

But what are we to conclude from these scenarios of theorising? Evidently not that theories are a ‘nice to have’, or an addition to the organism’s general functioning. Theorising is the *essential condition* of an organism in its attempt to survive and reproduce, beyond which it has no function. Some theories are useless or counter-productive. Some theories are

useful or effective. The brain cannot distinguish the useful from the useless beyond the physical fact that some enable survival, others do not. Even then, the brain cannot know one from the other as such: its ‘decision’ on use results from trial and error with feedback, or from ‘instruction’ from another. ‘Reasoning’ as a course of action establishment (active or not) – for example: the present is like the past (or more subtle means, as with Bayes’ theorem); or association (of thing/situation, or of feature, factor or flavour); or one and one is more than one (and so to mathematics) – must result from purely neural operations about world conditions. These will be signified as a series of brain-sign states erroneously designated as conscious cognitive processes (and thus specified here *as* e.g. ‘association’ or ‘mathematics’). The development of theories takes place over time from infancy according to the evolutionarily developed means of the brain (cf. Peirce & Husserl, and subsequent neurotheorists, e.g. Edelman).<sup>123</sup>

But, the question might be asked: How can the brain interpret its own causality as brain-sign (as e.g. association or mathematics) if that causality is dispersed within its neural network (+) conditions? If the organism is orientated causally as  $E=mc^2$  but (to us) invisibly so, how can the brain turn that dispersed orientation into  $E=mc^2$  as expressible as brain-sign? How does the unintelligible but causally effective brain turn into (or become neurally interpreted as) intelligible brain-sign? After all, for example, the speed of light ( $c$ ) is an *intelligible* absolute of the universe, is it not? To answer this question, we need to take a somewhat extended route, and in stages.

### **8.1.1 First stage: Current misrepresentation of the nature of the problem**

We return to the approach of Bechtel and Abrahamsen (cf. §4.3) in their book *Connectionism and the Mind*. (One might have hoped that Bechtel’s later book, *Mental Mechanisms* (2008), would be a development. Unfortunately it contains no index entry for consciousness or

---

<sup>123</sup> In her article, titled Inference to the Best Decision, Patricia Smith Churchland notes that for the brain, ‘case-based reasoning is more fundamental [than rule-based reasoning]’ (2009; 425). This is likely so. However, she consistently refers to the fact that reasoning often takes place unconsciously rather than consciously without explaining what the difference is or what consciousness is for. It seems more likely that the brain *never* operates by rule-based reasoning (because such ‘rules’ do not exist *per se*), but it ‘appears’ to as brain-sign.

explanation thereof, and confesses in the opening pages that what the mental *is* is unknown – which rather undermines the rationale promoted in the title.)<sup>124</sup>

I choose this text because, as its title suggests, there is supposed to be some direct relation between neural networks and the mind. According to the authors:

‘Although compromises may be necessary in the short term, cognitive models ultimately rest upon or constitute neural models; that is, they should specify the structure and activity of just those neural pathways that carry out the specific cognitive task being modeled.... Whether implicitly or explicitly, this position rests on a theoretical commitment known in philosophy as the *mind-brain identity theory*. In the most common version of this theory, type identity theory, particular types of mental states are identical to particular types of brain states’ (2002; 343).

The problems with this statement are immediately evident. Firstly, since the authors do not define mental states – although one must suppose they refer to seeing, thinking, feeling, etc., without (scientific) clarification of them – we are given no explanation of what mind-brain identity actually entails. We are already (although supposedly in the realm of science) in a fiction. Secondly, there must be the supposition that mental states are cognitive. But since we do not know what mental states are, their association with cognition is unexplained. What is not to be denied is that what is referred to as cognition, i.e. the ability of the organism to get around in the world successfully by its neural activity, depends upon the functioning of the brain (‘neural pathways’). Thus thirdly, what we are not given is an account of what biological method the brain uses to achieve its biological ‘aims’, and particularly how those aims (independently of notions associated with mental states, e.g. intentionality) are to be described. Cognition itself, therefore, is an entirely opaque concept. But the authors are so attached to this notion they continue, ‘Such an identity of mental and neural states permits an alignment between accounts – at the limit, a reduction of cognitive theories to biological ones’ (343-344). Thus this account self-defines its intrinsic limitations.

The authors contrast the methodology of identity theory with that of *functionalism*. As they describe it:

---

<sup>124</sup> ‘A serious obstacle to progress in...research is the lack of an understanding of what mental operations themselves are’ (2008; x).

‘For functionalists, mental states are characterized in terms of their interaction with other mental states and their relation to inputs and outputs. Although mental states can be implemented in the human brain, they can also be implemented in computers and other mediums. There is no systematic or necessary relation between mental states and brain states, and no point in trying to align cognitive and biological accounts’ (344).

Indeed, since functionalists often take the approach that the mind is the processing of symbols, which does not appear to be how the brain operates, there would be little point in functionalism aligning itself with the workings of the brain.

However, the same criticism of functionalism can be made as with the identity theory. Since functionalists never tell us what mental states are (in any ontological sense), apart from naming them – seeing, thinking, feeling, etc. – the notion they interact, or have inputs and outputs, is entirely vacuous (cf. §5.0).

The authors continue. ‘Most cognitive scientists have assumed some version of the identity theory in theory, but have been more or less functionalist in practice’ (*ibid.*).

Thus with no sense of irony or embarrassment, the authors propose that the entire topic – current neural science in fact, with the word ‘cognitive’ inserted in various locations – is built on the fiction that mental states exist *as which* cognition is to be associated or defined.

But there is a far deeper ontological problem which Bechtel and Abrahamsen do not address. How can there be an identity theory, or a theory of functionalism? How does a theory arise, and what job is it supposed to do? The presupposition, invisible in fact, is that theories just appear mentally, and theorists can conjecture on them in the refined ‘space’ of consciousness to which there is unfettered access. Their approach (entirely normal) is of a science *not* located in the human being as physical being. But *why*, quintessentially, is there more than one theory for the apparently particular circumstance? How could a conscious state not get things right with one theory by *being* knowledge?

### **8.1.2 Second stage: A current status in the Philosophy of Science**

Let us narrow the focus to science. We have emphasised the approach of science for brain-sign theory itself. Amongst other remarks we have said that: It was not until Newton’s discovery of gravity and laws of motion, that motion in the heavens conformed to a unifying

theory with motion on earth. The whole universe became one kind of thing (Chapter 1). In other words, science claims to establish the uniformity of the universe in an ontological sense by defining causal laws for entities in the universe. We have also said that: The key factor about physicality is not substantiality or materiality simplistically (or mechanically) conceived, but its behaviour under laws. The grasp that science gets on the physical universe derives from the lawfulness it finds there – signified by its representations of it (categories-of-the-world) (§7.2.2). This states that the scientific way to grasp the universe *as* physicality is to appreciate its conformity to laws intrinsic *to* physicality. But that does not mean that for us there is an easy access to the universe so that physicality can be so classified, or that we can inevitably define the universe as a whole *via* laws we can establish. I.e. brain-sign theory states that *our* laws are not necessarily nature's laws (if such exist), or that the entities we define as conforming to laws are entities in the universe.<sup>125</sup> Thus we need to contextualize these statements both in terms of scientific theory, and how science can be accomplished, given what we are.

Ronald Giere, in his article on scientific theories, states:

‘A model-based understanding of theories is easily incorporated into a general framework of naturalism in the philosophy of science. It is well-suited to a cognitive approach to science. Many forms of representation are presumed to be embodied as mental states in the brains of real people.... It is very plausible to suppose that physicists, for example, possess mental models, or at least partial mental models, for harmonic oscillators and other staples of both classical and contemporary physics. It is the possession of such mental models that makes it possible for them to recognise a new situation as one for which a particular sort of theoretical approach is appropriate’ (2000; 523).

Since Giere does not account for the mental states ‘in which’ the theories of scientists are deemed to be held and explorable (together with their functional characteristics – e.g. recognition, causality), we see that the foundation of science *as* theory is unexplained. Science as an enterprise is entirely unscientific, apparently!

However the development of model-based theories, as referred to, is highly significant. Giere contrasts these with ‘theories [as] well-defined entities’, the *classical* view

---

<sup>125</sup> Cf. e.g. Cartwright (1999).

as he terms it. The classical view is that theories can be understood to be a linkage of observational terms and theoretical terms.<sup>126</sup> Observational terms exist by direct access to the world to support their validity – e.g. we can see a billiard ball is spherical. Theoretical terms have only an indirect empirical validity, where an axiom system of the theory may have terms with empirical interpretation, as with Newton’s laws of motion. The obvious problem, brain-sign theory aside, is that while we can see the billiard ball is spherical, what of the shape of the moon or a proton? Without direct sensory access, our empiricism has become speculative.

By contrast, as Giere says of model-based theories,

‘What is needed is a weak notion of similarity [between the model and the real system], for which must be specified both in which respects the theoretical model and the real system are similar, and in what degree’ (*ibid.*).

In other words, the notion that theories can exist as well-defined (or true) entities (the classical view) is to be discarded. For as Giere states: ‘A model-based [or ‘semantic’] understanding of scientific theories provides resources for appreciating the illusiveness of theories in the practice of science’ (524).<sup>127</sup>

An example Giere gives of such a model (as in a previous quote) is for a simple harmonic oscillator, for example a weight suspended on a spring. ‘The state of the system at any instant of time is a point on a two-dimensional position-momentum space [a state space]. The history of the system is given by an ellipse’ (519).

The problematic question of how well the model (e.g. the ellipse) *can* match the ‘real-world’ (the ‘weak notion of similarity’) is indicated by Giere’s use of the word ‘illusiveness’. For, for example, as he says, equations associated with the harmonic oscillator, though scientifically useful in describing the system, are in fact for a perpetual motion machine, and such do not exist in the ‘real world’.

But what does the word *illusiveness* really point to? Is it merely a weakness in scientific method?

---

<sup>126</sup> Reference for this formulation is to Carnap (1956).

<sup>127</sup> Giere’s perspectivism is given in Giere (2006). The Introduction contains these two revealing sentences, presaging his thesis. ‘Like spatial perspectives, color perspectives are intersubjectively objective. That is, most people generally see the same objects as similarly colored in similar circumstances.’ The dependency upon consciousness, to which he refers but does not examine, contains a by now familiar unjustified assumption of its status.

Fundamentally (but unrecognised as such) it points to the *predicament* of the physical human being (Chapter 4). The predicament holds that there are no mental states in which theories exist as knowledge (*pace* Giere), which can be well-defined and causal. In other words, the underlying issue is not what science can achieve, but what the human being can accomplish as a physical entity *from which* science and scientific theories issue.

In fact the necessarily ‘weak notion of similarity’ is appropriately associated with the brain-sign account of theories, causal and communicative. The model-based approach in the philosophy of science has developed in-line with the theory of brain-sign even though it was not available. I.e. although the fallibility, or limitation of human laws of nature has become evident, the unfounded supposition that they exist as mental states remains without consideration of what mental states are, or whether they could do the job they are supposed to do (which anyway is not well-defined). The advantage of brain-sign theory for the model-based approach in the philosophy of science is that mentality can be discarded completely – *and replaced with a model-based approach to the human being itself.*

But brain-sign theory does not propose there is no representation. It redefines the function and ontology of representation.

### **8.1.3 Third stage: Uniting in review stages one and two**

To explore this further, let us return to the article by Markus Peschl to which we referred (§3.2.1.1). It closes with these words:

‘Scientific theories are not so much descriptions of an environmental phenomenon, but answers to how to deal with this phenomenon in the form of functionally fitting solutions’ (1999; 209).

This is correct, brain-sign theory would say, if the word ‘deal’ means: be orientated physically/actively in a requisite way towards entities in the environment. By eliminating mental states, this becomes a bio-functional progression beyond the (incomplete) account of model-based theories of science.

Contrast Peschl’s statement with Frederick Suppe’s. Suppe says that the most common use of scientific theories is as ‘a conceptual device for systematically characterizing the state-transition behavior of systems’ (2005; 1016).

Although Suppe's expression 'characterizing the state-transition behavior of systems' seems informative in describing what science is after in its theorising, he nonetheless fails to explain in any bio-functional way how the processes of science (as e.g. concepts) function in the universe, since doubtless he is, at least implicitly, as tied to mentalist suppositions as Giere. Indeed, Suppe's statement itself exists in a manner uncharacterised in the scientific process, but which must be so characterised to incorporate the statements of physical human beings into that process.

However, Peschl is unable to arrive at an adequate formulation of the required position. At one moment he states that: 'The dynamics of theories is realized in the dynamics occurring in the synaptic weight space and the genetic space' (1999; 208). That is, he attributes the causal power of the theory to the physical synapses of the brain – as we have proposed (roughly), and prior to that, as potentiality in the genetic foundation of the brain and nervous system. But at another moment he states that:

'Knowledge and/or theories become *tools* that are used for predicting, controlling, and manipulating the environment.... We use representational entities, such as concepts, symbols, language, etc., as means for manipulating and influencing the environmental dynamics and/or the (representational) dynamics of other cognitive systems...in physics (particles, waves, forces, fields, etc.), biology, or psychology (dynamics of propositions, "mind", etc.)' (*ibid.*).

But how can 'representational entities' or 'knowledge' be causal? There cannot be both synaptic causality as a scientific 'tool' *and* representation/knowledge as (e.g.) the concept *particle*. Unless, of course, by some mysterious means, as Bechtel and Abrahamsen propose, there could be an identity 'reduction of cognitive theories to biological ones'.

And indeed, as if to contradict his latter statement, Peschl determinedly proposes that 'The...representational structure does not have anything to do with an "objective" description of the environment' (208-209), which must be correct (and fits model-based notions) if we stick with the synaptic version of a theory – i.e. of a brain's causal physical states which cannot have access to the world *per se*. But what, then, is the point of the brain's representational structure which, according to Peschl, is apparently both knowledge *and* not objective? – an oxymoron.

\*

We should now note the difference from Churchland, who says: ‘The brain is a medium selected for its ability to assume, hold, and deploy the conceptual systems we call theories.... A theory is the brain’s way of making sense of the world in which it lives, an activity that is its original and primary function’ (§4.3).

We agree the brain holds theories (in two ways); but they are not ‘conceptual systems’ belonging to the domain of consciousness by which the brain ‘mak[es] sense of the world’. A person makes sense, not a brain. But such *persons* do not exist.

#### **8.1.4 The resolution of the science question**

Our discussion arose from the question: ‘How does the unintelligible but causally effective brain turn into...intelligible brain-sign?’

Our first answer is that we do not know how it creates its sharable representational structure from the brain’s causal structure. That is a topic for brain science (within the bounds of brain-sign theory). But we can state why it does, and we can clarify what the word ‘intelligible’ entails. In so doing, we reconstitute the previous discussion.

Just as the brain causes movement without conscious control, so the brain from its causal states interprets itself as communicable representations (ref. diagram in §6.5). These may be what the natural disposition terms ‘images’ and ‘discursive statements’ (and all the other features of brain-sign). In other words, for the super-organism, the brain has a capacity to render a communicable status – which can be scientific theories. Humans are not conscious individuals dreaming up theories they may share: rather, as organism, they *must* generate theories and they must communicate even if only implicitly *as* ‘images’.

But since the brain has no direct contact with the world, theories are limitless. The power of science, constraining limitlessness, is theory testing by experimental means. But even then, as is evident, many theories can satisfy the same experimental conditions. That is because, not least, the world elements to which the theories point *as* brain-sign are not actualities in the world, but constructs of the brain (categories-of-the-world) – a spherical billiard ball as much as a proton, albeit there is direct input about billiard balls.

Thus to say the brain creates the *intelligible* is to remain in the Kantian mode. For intelligibility relates to the mental subject, but there is no mental subject for scientific

theories. Indeed, there is no location (as in the notion of mental states) in which  $E=mc^2$  resides by which a scientist, or any other person, can know it.

Surely  $E=mc^2$  is a model-based theory. But it is a model *by* the neural brain, not *for* a mental subject. Scientists directly using the model are causally orientated toward the world in their neural states which signify by the model. But the mode as which scientists act in relation to the world is *not*  $E=mc^2$ . It is a dispersed physical property of their brains. We in whom the model also occurs, though we do not work with it, have it in varying detailed ways, for some are more scientifically orientated than others.

A vivid example of the two modes of theories appears with the *idiot savant*, or as now more customarily known, *autistic savant*. Dustin Hoffman, in the film *Rain Man*, plays a character (Raymond Babbitt) inspired by the life of Kim Peek, who performs astonishing feats of arithmetic and recall.<sup>128</sup> But that ability does not result from a capacity for ‘consciously’ and laboriously performing arithmetic sums with pencil and paper, as we normals do (supposedly). It is a neural ability of a particular brain, nonetheless communicated by brain-sign. When asked, in the film, how he gets to the answer to a calculation, Raymond says: ‘I see it’. I.e. *he* (person Babbitt) does nothing. The mystery of his capacities in the natural disposition is resolved by brain-sign function where consciousness does not exist.

(In mentalist discourse numbers are problematic – possibly abstract entities, possible theoretical, whatever those are deemed to be. However, numbers so conceived do not exist. In brain-sign terms, numbers are an aspect of language. They are communicative states derived from the action of the organism’s brain toward the world in numerical terms. It has been shown that some animals can discriminate quantity or ‘count’,<sup>129</sup> and as a category-of-interaction from the causal action, ‘numerosity’ must ‘hover over’ brain-sign objects.)

We have now clarified the Bechtel and Abrahamsen statements. They contrast identity theory models and functional models of the mind-brain condition, confessing eventually that ‘Most cognitive scientists have assumed some version of the identity theory in theory, but have been more or less functionalist in practice.’ We asked: But *why*, quintessentially, is there more than one theory for the apparently particular circumstance?

---

<sup>128</sup> United Artists, 1988

<sup>129</sup> E.g. Brannon and Terrace (1998). Contains references to other studies.

We have provided the rationale. As Peschl rightly says, in one breath: ‘Scientific theories are not so much descriptions of an environmental phenomenon, but answers to how to deal with this phenomenon in the form of functionally fitting solutions.’ Neither of the quoted theories (identity and functionalism) actually work, and apparently both are used at the same time which is a contradiction. We offer a third theory: brain-sign theory. The advantage of brain-sign theory (as with Einstein’s theories over Newton) is that it addresses more of the conditions of the situation – until and if others are brought to bear.

The bio-functionality of the scientific process can be now stated in reductive terms. Scientists pursue theories for control in the world (cf. §6.8). This is brain action in purely neural terms, inaccessible as brain-sign, although a signification of it as categories-of-interaction might be *sensed* (in the natural disposition) as ‘curiosity’, ‘concentration’, etc. But science is an essentially collective activity of the super-organism, and entails the continual (biological) attempt to dominate other brains by ‘our’ specific theory (cf. Kuhn §8.1.6.2). Scientists do not attempt to convince others of their view by credibility or logic: their brains attempt to program other brains in conformity (as language), a biological inevitability. ‘Credibility’ and ‘logic’ are brain-sign products without causality.

### **8.1.5 Error**

Although Descartes was quite aware of error, he attributed it to himself and not to God who was the guarantor of his knowledge. Specifically, he attributed it to his will which outruns his knowledge (Fourth Meditation). He also recognised his senses could deceive him.

Descartes’ view of what became termed consciousness was entirely bound up with God, together with the notion of faculties. But if what we see (for example) is supposed to correspond to what is out there independently of God, then (as we have discussed at §3.1.3) the integrity of the notion is put under strain if we see something other than what is there, for example a lamppost instead of a tree. How could the mind misrepresent? Slezak’s way out of this is to dismiss error because error can only exist in the view of the external observer, who in fact does not exist. The problem with this approach is that it demolishes the notion of consciousness altogether.

With brain-sign theory, however, the topic of error is recast. For there need be no attribution of truth or correspondence between world and representation because the brain

does not ‘replace’ the world with its representational states of mind upon which it then causally operates. Representation is a signification of the cause of its causality. The brain does not know why it acts, and it does not represent that knowledge.<sup>130</sup> It *signifies* why it acts, and the signification is an artifice which may, to an adequate degree, function in the unit of the communicated.

Thus the reason we may ‘see’ a lamppost rather than a tree is not because our mind misrepresents, but because our neural status is orientated toward the object ‘out there’ as a lamppost because that is how it has been interpreted by the brain. And that is because most of what causes anything to be responded to *by* the brain already exists *in* the brain, and is triggered by limited input. Pechl puts this well.

‘As most neural systems have a *recurrent architecture*, the internal representation state is determined not only by the current environmental input (which is supposed to be represented in the traditional view), but also by the previous internal state. The current input can only select from a set/space of possible successor states. However, this set of successor states is predetermined by the neural architecture and by the current internal states. So, there is no way of guaranteeing a stable referential relationship between repraesentandum and repraesent’ (1999; 196).

Brain-sign theory expands this explanation because brain-sign representation makes no claim to be accurate or objective as a *mental* operation. Thus it avoids Pechl’s unresolved dichotomy between the neural and the representational as causal.

### 8.1.6 Adjunct topics

For completeness, we should remark briefly on debates in the philosophy of science illuminated by brain-sign theory. We have already laid the foundations.

---

<sup>130</sup> How the brain operates causally without mental content is a topic under current discussion. See e.g. Peter Machamer, particularly 2009; 173. “Suppose the atomic elements (individual axon or dendrite processes, say) don’t have content but respond differently to different inputs connected directly or indirectly to environmental influences...” etc.

### 8.1.6.1 Realism/Instrumentalism

The key word here is *reality*. Presumably the universe is real, and reality is physical. The question is, how do we get in touch with it? The mentalist proposes the mental is also real and is the means. But this sets up the problematic conditions of the mental's nature.

Philosophers of science seldom explore the mental *qua* mental: indeed, the default supposition for empiricism is that mental states *provide* the mechanisms for 'getting in touch'. Hence Giere: 'It is very plausible to suppose that physicists, for example, possess mental models', and thence our comment: Observational terms exist by direct access to the world to support their validity – e.g. we can see a billiard ball is spherical (§8.1.2). But at the same time there is Bechtel's deflationary statement: 'A serious obstacle to progress in...research is the lack of an understanding of what mental operations themselves are' (§8.1.1f).

However, once brain-sign theory and its critique of mentalism is grasped, Jarrett Leplin's statement that 'The debate between realism and instrumentalism is at an impasse' (2000; 393) is seen as inevitable, for the very *terms* 'realism' and 'instrumentalism' result from an erroneous ontology.

To propose mental states are *real* requires explanation of what they are. And what they are requires them to be located in the physical universe. This has proved impossible. In fact, therefore, the term *mental state* denotes nothing at all beyond a (displaced) category-of-the-world as brain-sign (§6.12.3).

Of course, to say they are not in the physical universe results from the theory of brain-sign. The legitimacy of brain-sign theory (in the natural disposition) lies in its explanatory power. But even that (the theory itself) is 'merely' a brain-sign occurrence (in brain-sign theory). The theory states that from the causal orientation of the brain toward the world, brain-sign results. However, once this stage of explanation is reached (i.e. we accept brain-sign theory as the *ultima ratio*), we have obtained the *grounds* for disabling the notions of realism and instrumentalism.

Which is that there are no scientist-mental-subjects holding theories, either as empirically *real* sensory states, or as (mental) *instrumentalist* states whose posited entities 'merely' function to render the theory workable. For the brain is already a physical

instrument which enables the human organism to function in the world, not a separate kind of being with indescribable (mental) access to the universe *per se*.

Thus all brain-sign states are instrumentalist, if one cares to use the term. But they are not causal and do not effect science. They effect communication between (amongst others) scientists about the collective scientific practice.

### **8.1.6.2 Thomas Kuhn**

In his *Structure of Scientific Revolutions* (original publication 1962), Kuhn proposed that science was not a continuous progression of increased rational grasp of a world objectively 'obtained' in a neutral observation language, but rather shifts between paradigms of radically different underlying hypotheses about the world, and could not be regarded as necessarily progressive. His position has been termed, therefore, relativistic (with which he did not concur). A representative passage is as follows:

'Successive paradigms tell us different things about the population of the universe and about that population's behavior.... But paradigms differ in more than substance, for they are directed not only to nature but also back upon the science that produced them.... As a result, the reception of a new paradigm often necessitates a redefinition of the corresponding science.... As the problems change, so, often, does the standard that distinguishes a real scientific solution from a mere metaphysical speculation, word game, or mathematical play. The normal-scientific tradition that emerges from a science revolution is not only incompatible but often actually incommensurable with that which has gone before' (1996; 103).

Kuhn's thesis was obviously not worked out, as brain-sign theory, from brain function. Thus from the present perspective, Kuhn's engagement with how human beings create different theories, though pregnant, is incomplete. *Structure* contains no index entry for consciousness, nor reference to the suppositions it entails. Yet (in the postscript added in 1969) he does identify relativism in the physiology of the human being.

'Much neural processing takes place between the receipt of a stimulus and the awareness of a sensation. Among the few things we can know with assurance are: that very different stimuli can produce the same sensation; that the same stimulus can

produce very different sensations; and, finally, that the route from stimulus to sensation is in part conditioned by education.... Notice now that two groups, the members of which have systematically different sensations on receipt of the same stimuli, do *in some sense* live in different worlds' (193).

In other words, Kuhn casts doubt upon direct sensory input for validating scientific models objectively because: (1) we do not necessarily see the same thing from the same input; and (2) what we see from the same input can be different depending upon the influence of our education upon what we see. That is, we are conditioned by our history and thus see 'the same' *via* that conditioning rather than as a result of 'unvarnished' input.

But his approach to neural activity becomes confused when pushed further.

'Once we have learned to do it, [perceptual] recognition of similarity must be as fully systematic as the beating of our hearts. But that very parallel suggests that recognition may also be involuntary, a process over which we have no control [because it is a result of neural processing]' (194).

Brain-sign theory would concur with this, of course. But Kuhn continues:

'After we have had a sensation, perceived something, then we do often seek criteria and put them to use...in interpretation, a deliberative process by which we choose among alternatives as we do not in perception itself' (*ibid.*).

And this leads him into a mentalism neither accounted for nor justified.

'We try...to interpret sensations already to hand, to analyze what is for us the given. However we do that, the processes involved must ultimately be neural, and they are therefore governed by the same *physico-chemical* laws that govern perception on the one hand and the beating of our hearts on the other. But the fact that the system obeys the same laws in all three cases provides no reason to suppose that our neural apparatus is programmed to operate in the same way in interpretation as in perception or in either as in the beating of our hearts' (195).

Kuhn's 'involuntary' nature of perception more properly indicates there is *no* perception, as brain-sign theory states. This is then of a piece with the notion that deliberation is also involuntary, indeed *not* deliberation. 'Physico-chemical' laws are inescapable. Yet

Kuhn held onto a mentalism at the very threshold of what would have been an entirely neural analysis, that which leads to brain-sign theory.

Indeed, Kuhn's insights are greatly assisted by brain-sign theory, for it clarifies in physical terms how the 'different worlds' in which we live, and as which scientific theories arise, are to be accounted.

Later, in 1993, Kuhn's reflections on his lifetime's work offer interesting parallels with brain-sign theory (though he never escaped mentalism, or the supposition that language has meaning). For example:

'Understanding the process of evolution has in recent years seemed increasingly to require conceiving the gene pool, not as the mere aggregate of the genes of individual organisms, but as itself a sort of individual of which the members of the species are parts. I am persuaded that this example contains important clues to the sense in which science is intrinsically a community activity' (1993; 329).

Had he developed this idea ('a sort of individual [made up of] members of the species') and dropped mentalism completely, he could have alighted on the notion of the super-organism. The 'clues' would then be made concrete in brain-sign theory.

And again, concerning the relation of biology to science.

'The development of human culture, including that of the sciences, has been characterized since the beginning of history by a vast and still accelerating proliferation of specialities. That pattern is apparently prerequisite to the continuing development of scientific knowledge.... What permits the closer and closer match between a specialized practice and its world is much the same as what permits the closer and closer adaptation of a species to its biological niche. Like a practice and its world, a species and its niche are interdefined' (336-337).

Brain-sign theory explicitly proposes that science, and scientific theories, are biological adaptations (a modality of potential biological 'fitness'), and do not rise 'upwards' into the disembodied strata of consciousness.

## **8.2 Afterword**

As was said at the beginning of the chapter, we can only glance at some indicative examples. Disciplines are founded (usually implicitly) on the notion of consciousness, with the ensuing knowledge it (supposedly) confers.

Psychology, religion, sociology, law, ethics, medicine, philosophy, art, anthropology, evolutionary biology, politics, conflict resolution, history, cultural studies, literature, education, linguistics,...

The theory of brain-sign, by altering our assessment of what a human being is, will alter all disciplines fundamentally.

## The Theory of Brain-Sign

### 9.0 Considering ourselves

By demonstrating the mind is not the brain, indeed that the mind does not exist, we have altered the entire context of our 'self-understanding'. Consciousness specifically is a brain theory of itself. But as we have seen from brain-sign theory, that means that 'we' *as* consciousness (as the brain phenomenon so-called) are the brain's self-explanation for the purposes of communication. Since neither mind nor consciousness exist (brain-sign theory claims), neither gives us a hold on the universe as the natural disposition supposes (the megalomania to which we have referred, §7.8.1).

The question now is: What impact does brain-sign theory have upon our own state as brain-sign? For whilst our natural disposition cannot be wholly dispelled – just as our position on the earth retains the sense of 'up there' and 'down here' – our redefinition by the brain as brain-sign introduces a self-interpretation *in* that state which is entirely different, *viz.* the scientific disposition.

\*

*Preamble: The conceptual metamorphosis*

The reader will have encountered the difficulty in reading the text, that there is the constant need to reposition our self-orientation from that of 'subject *reading* the text' to that of 'biological object as which the text appears' – and that for others, though probably they are not present. And to realise: (1) Not only that this is not a process of choice but an effect in our own brain; and (2) Despite the permanence 'the object (i.e. the text) of us as object (brain-sign)' seems to purport, it will never occur again, for it is a transient state of our brain.

'Becoming an object' is not taking a scientific attitude toward one's existence (biology), but rather towards one's so-called inner state (neurobiology). We are not in conscious command of ourselves in picking up a text and reading it. We are wholly determined in the process by our brain beyond any awareness of our self, or conscious decision. Wittgenstein pointed to an analytical difficulty concerning his own approach.

'How does the philosophical problem about mental processes and states and about behaviourism arise? – The first step is the one that altogether escapes notice. We talk of

processes and states and leave their nature undecided. Sometime perhaps we shall know more about them – we think. But that is just what commits us to a particular way of looking at the matter. For we have a definite concept of what it means to learn to know a process better. (The decisive movement in the conjuring trick has been made, and it was the very one we thought quite innocent.) – And now the analogy which was to make us understand our thoughts falls to pieces’ (1967; para. 308).

Brain-sign theory does not perform the first step unnoticed – assuming mental states exist. It inserts itself before the first step to question its legitimacy and declares it wanting. But the result is not something we then know in another way; rather it is an alteration in how we ‘find’ ourselves, i.e. we *are* different processes *as* neural communication.

In Chapter 1, comparison for addressing the problem of mentalism was made with grasping that the earth is not the centre of the universe, that one’s subject position on the earth must be overcome to locate our cosmological status in a scientific way – whence our position becomes that of (cosmological) object. In a way Psychology as a discipline had achieved a step in this process by proposing that our experience does not derive from our spontaneous subjectivity (or quasi-divinity), but conforms to laws to which we are inherently conditioned. This ‘fact’ is what, in principle, gives Psychology (*folk psychology*, as the literature has it) its scientific status: it allows us, in self-consideration, to objectify ourselves by *having to be* just perceptions, thoughts, feelings, etc. The problem (which Husserl partially saw) is that Psychology (as a false step) does not escape common-speak, in our case (and by contrast with Husserl) the notion of the subject and its experience. It does *not* identify objectivity, as does brain-sign theory, which therefore rejects perceptions, thoughts, feelings (of a subject).

An immediate example is reading. We suppose, by some unknown means, that information is sucked up into our heads as meaning that we understand. In our heads we suppose there is a vast reservoir of previously established literacy into which our current reading fits. But as demonstrated (Chapter 7), this miscasts the process, thus disguising our nature. Of course it is easier and more appealing than brain-sign theory to assume this sucking up, meaning and understanding of an inner mental life which *is* us – though it makes no physical sense. But our aim is not to endorse the common assumption (which we do in ordinary life), but to expose our physical actuality. Then we will have a normal science in a Kuhnian sense.

In turning to our ‘self-consideration’, there is an even greater need (than in cosmology) to ‘attend to’ the processes we are. We must allow the revision to play out in us, as which the processes will self-reveal *as us* though not *to us*. As in the previous chapter, given the vast potential scale of the analysis, we can only give a few representative examples.

## 9.1 Light

For the natural disposition light not only allows us to see, thus giving direct access to the world, it is associated with life (being alive in the world) and truth (that which can truly appear) (§2.6).

Brain-sign theory rejects this. Light ‘seen’ is a communicative representation without impact on the causal properties of the organism. The brain’s operation in the world requires the presence of electromagnetic radiation; but this entails, from the eye as input medium, the impact of transduced signals on the physical-chemical processes of the brain. Light ‘seen’ is not electromagnetic radiation (*contra* Churchland, §4.3). It is an inter-organism brain-sign for the unit of the communicated. That life or truth results from the presence of light *in us* is false.

Indeed, in brain-sign terms, the word ‘life’ is a suspect term. If we suppose life depends upon the capacity for being conscious, then we never die because we are never alive (cf. §2.4). The term ‘life’ must be used in a more precise biological sense.

The natural disposition forces our seeing to be our prime and immediate grasp of the world. Overcoming this greatly assists command of brain-sign theory. The *overcoming* is the brain altering its physical stance to the world; the brain-sign result will arrive necessarily. It requires practice which is, of course, the brain altering its structure (without brain-sign influence).

## 9.2 Our reality (by contrast with mentalism)

Reality is a term we have mentioned in various contexts. In common parlance it is without precision for, *per se*, it is unidentifiable. Only with the scientific disposition can it be identified.

The common parlance problem is that we are confused by our condition (as Husserl said), yet that does not stop us functioning. If it did, perhaps we would be forced to be accurate. Nature, however, is not concerned with scientific theorising, for the function of our existence is to survive and reproduce, not develop scientific theories. By evolutionary chance and behavioural modification (and beneficially), scientific theories have emerged, particularly over the last four hundred years.<sup>131</sup> With brain-sign theory, the neural phenomenon (misunderstood as consciousness) is part of a wider (to be developed) brain science.

The *causal reality* of our biology is the physical properties of us as organism, and the forward move (as it were) of our existence is determined by those properties. This is unknown to and unknowable by us for two reasons. The first is that we cannot know anything as a result of our biological nature – i.e. our *significatory reality* as brain-sign is not knowledge. The second is that, insofar as our brains interact with the world and theories arise as brain-sign from that interaction, we must suppose the functional complexity of the brain outstrips our (i.e. the brain's itself) capacity for analysis of it. (We commenced reference to this at §4.2.)

Thus, if it occurs to us, we may say out loud or not: 'I am picking up my cup because I want to drink my tea.' But the cause of our picking up the cup is not the desire to drink the tea: the cause is an unknown. In other words, nothing occurs *to us*; rather it occurs *as us* (cf. §6.1). In a similar way, the lion sets off on its chase of the gazelle. But it cannot say: 'I am setting off on the chase because I want to catch that animal and eat it to alleviate my hunger.' It knows not that it is hungry or that it sets off on the chase to eat the gazelle. It is our deception (which the lion has not) that our knowledge is the cause of our actions.

But regardless of our technical limitations *vis-à-vis* brain analysis, for ourselves as brain-sign we have an identification of the cause of our state which with mentality was entirely lacking (an identification which is part of our function as a neural sign). Thus although that cause is *per se* unreachable by us *as* brain-sign, there is nonetheless an improvement in our 'self-relation' as organism (cf. §6.12.3) because identification occurs in, and is of the physical world. Here are two examples.

---

<sup>131</sup> This is not to imply that there was no such thing as science before the last four hundred years. It is a question of scale and systematicity.

1) In the natural disposition our mental life is our reality, and it concerns us. We wish to be happy, not sad; successful (in whatever terms), not unsuccessful; feel well, not ill; appear honourable, not dishonourable; and so on. These wishes, and numerous others, can seem to result from goals in the world which, when realised, will cause our hoped for mental state. Indeed, so encased are we *as* mental states, we do not differentiate them from the physical world *per se*. The physical world of our goals is (we suppose) the one we see, hear and feel something about.

Although our certitude seems unassailable it is entirely false, and our so-called experience points to this: we do not find our hoped for mental state necessarily follows from what is realised. The pointing-to of our 'experience' is ineffective because it is not underpinned by valid theory. Indeed in ontological terms, how our hoped for mental state could arise from our relation to the physical world is a complete mystery.

By contrast, in the scientific disposition we see that what is significant about us – our physical status in the world that causes us to act – is not determined by us as brain-sign. Therefore we cannot make ourselves happy by altering our mental states (including those supposedly betokening the physical world) *as if they were part of the (causal) physical world* because we have no, or are not, mental states. Happiness definitionally does not exist: it is not a physical property.

The *sense* that we are happy derives from the immediate state of our casual brain. The signification of 'happiness' as a category-of-interaction results from our organism's causal relation to the world. But this is not a world known; and our causal condition does not indicate a rational or indeed irrational assessment of our relation to the world on a knowing principle. The cause of the sense of happiness can only be specified by the physical-chemical conditions of the neural brain. Without specification by a brain science, the sense of happiness is inexplicable. Even with a brain science, that sense is a moveable feast both in terms of what causes it, and in the nature and degree of how it is 'felt'.

Now although it might seem our condition as brain-sign is involuntary – almost one might say inconsequential – this is a misreading: indeed, the 'seeming' is a hangover from the natural disposition. Just as our causal relation to the world results from our causal brain, so our brain-sign states result from our brain's causal conditions, and to alter our brain-sign states requires us, *as* causal brain, to do the apposite things. Thus in 'self-perception' we have

to adjust, not to the notion of involuntariness, but to our physical being as our causality which occurs before our resulting state as brain-sign (i.e. *not* as awareness).

Of course we cannot know how we as brain-sign arise in a technical sense, any more than supposed mental states; but on the other hand (as said) we are no longer in misconception about our causal physical nature, nor how we as brain-sign arise in a functional sense – which is a vast improvement (adaptively) from the obscurity and mystification of mental states (again, cf. §6.12.3).<sup>132</sup>

In practice we should not concern ourselves about our mental states, because there is nothing we (as mental states) can do about them: they do not exist. More accurately, an adapted brain operates without generating (e.g.) undue ‘hope’ or ‘fear’.

2) Consider the example at §3.3.4.2 of the paralysed subject who appeared able to move a cursor on a screen by their thoughts. Since the subject has no thoughts, they could not ‘move the cursor’ by them. On the other hand, the ‘self’-training involved in moving the cursor, while thoughtless, nonetheless entails the subject’s brain organising itself in its causal processes to send motor signals that will move the cursor. And why not, since that same brain could in principle send motor signals to the hand, etc., as the writers say. Thus the neuroscientists, and Clark and Prinz, simply misinterpret what the experiment demonstrates because they make their deduction on the wrong model. The subject may brain-sign signify with what they take to be sight, thought and will, but ‘what they take to be’ is the natural disposition in operation.

In relation to the world, *all of us* are ‘paralysed’ in terms of brain-sign. But that does not prevent us being the causality we are as our brains. Adjusting to this scenario is a means of identifying reality in the nature of our being.

### 9.3 Transparency; opacity

For bio-physical communication, we are evolved to be transparent to each other. *We inhabit* a common world by our projection *of it* in terms of our causality in relation *to it*. We are not

---

<sup>132</sup> Churchland’s notion (1981) that there *could* be a ‘completed neuroscience’ is not based upon the feasibilities of neuroscience because his account (as already explored) presupposes a knowledge of it which he has not explained.

individual subjects for whom the (common) world is present and thence upon which we act (the megalomania).

When we are part of thousands at (e.g.) a football match, the match is not ‘out there’ which we all view from ‘in here’. Our participation is a neural orientation (each of course individual), and our ability to share ‘the experience’ (as organism) results from the commonality and difference of our state as brain-sign.

Since this scientific disposition viewpoint does not appear immediately to alter how we engage with the football match, or share it with others, it might seem a purely abstract formulation, adding nothing to our actual experience. But the change in viewpoint will have a profound influence on us as it is neurally absorbed.

In the natural disposition our enjoyment of the match (or not) appears to *be* our mental condition, which fits into our mental life generally. It results from the mental interaction we have with the match. We do not necessarily consider that interaction *as* mental; it is just ‘how things are’, mentality being an implicit (cultural) assumption. But there can be no implicit assumptions with brain-sign.

Brain-sign theory demonstrates our relation to the match is purely neural. We ‘enjoy’ the game because it stimulates our neural structure signified as the enjoyment we ‘feel’. This category-of-interaction accompanies the fact that we ‘see’ the game (as categories-of-the-world), but it is not caused by it.

Our exclamation with our companions upon a goal scored does not result from seeing it, but from the common source which is the stimulated neural structure in each of us. We *seem* to enjoy the goal scored because our team is winning and performing well. And while we are embedded in the natural disposition, that we enjoy our team’s performance, indeed that we support a team performing well, is also natural.

But the scientific disposition shatters the impenetrable *façade* of mental states. Once observers of the game and our own enjoyment, an orientation has developed *as us* that we do not observe at all. Our physical relation to the game is manufactured by our neural brain, and essentially arbitrary though determined.

In deflating the notion of mentality we appreciate that *it does not matter* to our brain that we ‘enjoy’ football and our team winning: ‘having’ a team is meaningless. ‘We’ are the mere locus of neural events, genetically prepared and historically manufactured through time.

It does not take brain-sign theory for us to imagine that everything about ourselves could have been different – at least within our given physical structure. Brain-sign theory further demonstrates that the neural communication brain-sign effects *hides* the nature of our causal existence in fact. Peirce and Wittgenstein (for example) supposed the brain phenomenon to be our selves for, despite their revisions of mentalism, they took consciousness for granted;<sup>133</sup> but our causal being has an entirely different order of existence which is invisible (or inaccessible) to the brain phenomenon as brain-sign. Worse still for mentalism, what is apparently ‘revealed’ as ourselves as brain-sign results from a causality that not only outstrips our capacity to grasp (for brain-sign is not a grasping), it forever precedes what we are *as* ‘revealed’. Our supposed self-knowledge, insofar as it is a consistency (that we enjoy football matches and support this particular team), is rather the repetition we have become from our causal neural states. If that ‘self-knowledge’ is to change (i.e. if ‘we’ are to change), it will result from that inaccessible source, i.e. the neural brain. It will be a new repetition, *not* self-knowledge.

Thus while brain-sign appears to be a revelation about ourselves *to* ourselves (as in the metaphor of illumination for consciousness), it covers over a causal opacity/inaccessibility which lies at the heart of our being. The revelation is not *for* us.

#### **9.4 Heaviness; lightness**

We designate objects as heavy or light. An elephant is heavy; a feather is light. Curiously, we seem to experience ourselves in an analogous way. Sometimes we feel heavy, and sometimes we feel light. But of course the brain has no knowledge of heaviness or lightness about things in the world or its own states. The brain is simply a physical object, whereas heaviness or lightness appear to require a mind – one which can render analogies of sense. The ontological

---

<sup>133</sup> Wittgenstein appeared to suppose his later language theorising avoided the problem of consciousness, but it failed without his realising it. Of the *Philosophical Investigations*, etc., Bertrand Russell said: ‘The earlier Wittgenstein was a man addicted to passionately intense thinking, profoundly aware of difficult problems of which I, like him, felt the importance, and possessed (or at least so I thought) of true philosophical genius. The later Wittgenstein, on the contrary, seems to have grown tired and to have invented a doctrine which would make such activity unnecessary.’  
Quoted from Ray Monk (1990; 472).

muddle caused by mentalism is evident here: in the natural disposition we pass straight over it.

For brain-sign theory, however, the relation of brain to world is exclusively in terms of action. The supposed weight of an object relates to our body. We approach it as to be acted upon. We do not know this, of course, for we do not know anything. But in the super-organism our brain needs communicate upon our neurally activated physicality. Thus the *sense* of heaviness or lightness is a neurally invented category-of-interaction as brain-sign which signifies our *neural condition* in relation to the object (category-of-the-world).

But why can we apparently be orientated towards *ourselves* as heavy or light? What is the analogy? For mentalism this is inexplicable. We are no more physically heavy now than the lightness we were yesterday. How, on the other hand, could our (mental) mood be classified as heavy or light: how can a mood *be* heavy or light?

If we remove knowledge from the universe, the situation begins to clear. The senses of heaviness and lightness are purely neural significations. The elephant is not heavy and the feather is not light – which are mentalist suppositions. The error lies in supposing a replicative correspondence between the designations heavy and light as knowledge *in the world*, of objects in the world. If, on the other hand, the senses of heaviness and lightness are neural models for inter-organism signification (§8.1.x), then whatever words we use for them (e.g. in whatever language), and whatever our individual versions of them are *as* sense, they do not aspire to conditions of world replication (or representation) as knowledge, i.e. *as* in the world. They are completely localised within brains as neural signs to enable collective action.

Concerning our mood, we might have supposed that light things in the world are easy to move; heavy things difficult. By analogy, we might have supposed heaviness is the difficulty of moving ourselves; lightness an easiness. But this cannot be the explanation, for it would still involve an analogy with the world, and brain-sign cannot be such an analogy.

But did we not say (§8.2.3)<sup>134</sup> that depression might be supposed a heaviness, being related to the inability to act, an inability to move ourselves? The answer is: No. For that would allow a mentalist supposition to creep in. What is signified by so-called depression is a *neural condition* of the inability to act. Quote: ‘So-called depression is the *result* of an

---

<sup>134</sup> 8.2.3 is one of the sections removed to meet the length requirements of the PhD. However, the point is quite understandable here.

inability to act, to have control in neural terms.’ It is the neural terms that signify by so-called depression, i.e. the control system *for* action. There is no analogy with weightedness in the world, for the weightedness we ‘feel’ as depression is a signification of our neural condition.

But our so-called heaviness may be unrelated to so-called depression, i.e. as simply the neural condition of the inability to act. Creative individuals regularly talk of the heaviness or immobility of their state before a sudden outburst of creativity. What seems to be indicated is, not that the individual is depressed (as an analogy of weightedness); rather that the brain is preoccupied with its own restructuring (which is not apparent in any brain-sign terms) which subsequently will emanate as a new and interesting neural orientation signified as brain-sign.<sup>135</sup> Indeed, that emanation may be accompanied by, or result in, a ‘feeling’ of lightness and fluidity. In other words, although there may be an *apparent* ‘feeling’ of heaviness beforehand, it again (as with so-called depression) requires a *neurobiological* interpretation signifying something entirely different (i.e. the condition of neural states themselves), not a mis-analogy with the heaviness of objects in the world. (This is but one example.)

On the other hand, a ‘feeling’ of lightness cannot be taken necessarily as a beneficent sign of freedom or mobility, however ‘pleasing’ while present. Whatever neural processes are taking place, the actual state of the organism in the universe may be deeply threatened (etc.), which does not correlate with the ‘feeling’: it is, in reality, delusional (about the world). Such a case is cited by Damasio of patient S.

‘It was as if negative emotions such as fear and anger had been removed from her affective vocabulary, allowing the positive emotions to dominate her life’ (1999; 65).

This ‘clinical condition’, Damasio states, resulted from brain damage to the amygdalae on both sides of the brain. But one hardly needs to cite a clinical condition to appreciate the sense of ‘false optimism’ to which we are prone – as one example of the ‘irrational’ (i.e. inadequate neural-worldly orientation) state of lightness.

That the brain signifies heaviness and lightness in the world that *seems* paralleled by its own mental states is a ‘design’ requiring neural, not mental interpretation. I.e. we can be entirely misled by what the brain appears to connote by these states in mind terms: neural interpretation reveals precipitating conditions which are *not* analogous to world situations.

---

<sup>135</sup> E.g. of Freud, Ernest Jones stated: ‘Increased discomfort, with various symptoms of general malaise, always preceded Freud’s best work’ (1964; 460).

(For ‘mood’, the categories-of-interaction relate to the brain states themselves as categories-of-the-world, which is why they are non-specific, cf. §2.11.)

These examples demonstrate that retaining mentalist terms for brain-sign states is not apposite, but as said (§6.12.7), we cannot develop that further here.

## 9.5 The absence of memory

We should note that in neuroscience, memory is not regarded as a fixed property of the brain. Edelman and Tononi say, for example: ‘It is [the] property of degeneracy in neural circuits that allows for changes in particular memories as new experiences and changes in context occur. Memory in a degenerative selectional system is recategorical, not strictly replicative’ (2000; 98). So it is (hypothetically) that when we remember, our memories are qualified by our subsequent history.

Edelman’s problem, however, is that he fails to explain consciousness. For he continues: ‘In higher organisms every act of perception is, to some degree, an act of imagination. Biological memory is thus creative and not strictly replicative’ (101). (Kant said almost precisely the same thing, §2.10.) The question is: What is memory *per se*, or imagination *per se*, and how do they *per se* function in the organism? (*Vide* §6.10.)

Brain-sign theory explains (*pace* Edelman), not a changeable mentalist memory, but the constantly adapting causal orientation of the organism to the world. From this ‘the past’ occurs as brain-sign (repetition) – and is thus changeable.

## 9.6 What reflection is actually

The fundamental error of mentalism, upon which the entire notion depends, is that of reflection. As quoted in Chapter 1, Block says: ‘The starting point for work on consciousness is introspection and we would be foolish to ignore it.’ But the notion of introspection, or reflection, requires a physical model for mental states. Block offers none, and none is forthcoming generally from those who regard mental states as real.

Even so it seems we can interrogate our own minds. For example, if we forget a name, we struggle to look inwards to force the name to come to us. If we are trying to

remember directions, we try to visualise the environment and situation of the requisite locale. If we wonder what we think about some occurrence – e.g. do we approve or disapprove – we seem to look inwards for our immediate or considered response. There seems a lot inside us just as available as the world we see (which we do not, in fact).

Since this thesis has been directed to the alternative to mental states, there would be little point of addressing the topic all over again. But one essential element (and one example) can add clarity from brain-sign theory.

We *suppose* we try to think of a name we have ‘forgotten’. This is how we take it to be in the natural disposition, and the brain-sign occurrence appears no different.

But *of course* the brain-sign structure is as ‘we take it’ in the natural disposition because that is the way the brain structures what is conveyed as brain-sign. (E.g. brain-sign ‘appears’ as seeing, feeling, thinking which the brain effects without seeing, feeling and thinking.) Thus the cause of the causality (brain-sign) is signified as a self in its inner quest for the forgotten name which lurks within the mist of forgetfulness. And we ‘feel’ the effort of our internal probing. But no self exists in the brain, and it cannot search in the highways and byways of neural structure to spot the word lurking behind some neural hedge. The brain-sign ‘model’ of trying to find a word is couched in the fabric of interaction with the external world. But that supposed external world is also devised by the brain to signify (customarily) what has caused a causal neural state in relation to the world – a world which does not exist *per se*. In other words, *we do not experience* trying to remember a word. We are simply being the neural operation of the brain. (Why else would the brain fabricate such a scenario *of itself* unless it required to communicate its activity to others?)

But the brain can *add* to that operation (i.e. to our natural interpretation) as the scientific disposition. In so doing it can express the neural model of itself as brain-sign which does not suppose brain-sign is a mode of reflection (or introspection). It can express it in physical terms not proposing a self on an inner quest for words lurking in the brain. Here is that account.

Either by self-activation or prompted by external events, the brain attempts to brain-sign a name. For example, I may wish to tell you of a restaurant in town which is the perfect place to go. But I can’t remember it, or more accurately, my brain in its activity as the unit of the communicated cannot alter your brain causally, *via* compression waves, with what will

brain-sign in each of us as the name of the restaurant, because its causal states will not orientate correctly to that physical place and thus project associatively the name (even though an *'image'* may arise – ‘I can *see* it! I just can’t think of the name.’ ).

But my brain *must* activate the causal structure that will brain-sign as the name. The process of attempted self-activation is what the brain signifies as myself (my body in fact, §6.9) searching in my memory. What is evident is that the more the brain communicates this explanation of its activity (as brain-sign), the less able it is to perform the task of activating the appropriate causal structure. So it drops the self inner searching, and signifies as ‘going blank’, in which state it will more likely self-activate successfully. Indeed, the brain may activate other causal networks of association to try to get linked up to the precise causal network that will signify as the name. Eventually (if I’m lucky) the name will suddenly appear, and I will blurt it out, with a (brain-sign) sense of relief. Yet have no ‘idea’ how it happened.

Of course the process of self searching seems to be what is most personal and quintessentially me. There I am doing it in my mind. But that sense of me-ness in the process is brain-signed for others: nothing remains over as (knowing) self-identification to me.

## 9.7 The brain is an alien

The brain litters human history with erroneous suppositions, and does so *in us* every minute of the day (and night). As brain-sign, we are the hazard of its biological creativity to effect communication with others about the world of causality (particularly itself) of which it knows nothing.<sup>136</sup>

In our organism’s encounter with the world we are continually at risk. The proposition of evolutionary theory, survival and reproduction, is that we continue being and reproduce because we must, evading the risks. But this explanation cannot be knowledge. It is the product of a brain as communication about the cause of its causality in relation to the world. The brain continually keeps the organism going, and thus it hypothesises that the cause of *that* causality is a ‘force’ or condition of pure physicality organised in a certain way. It is an account of its own modus in the world described in terms of the world, a world of which it

---

<sup>136</sup> That is the case with brain-sign theory itself. But it is evidently a better theory in the physical universe than consciousness, though we do not know this – it is ‘merely’ how we are made operative.

knows nothing. The brain hypothesises as brain-sign, not so it can act upon its hypotheses, but to share with others scientifically. Its actions remain where they always were: its physical/causal conditions (a different hypothesis-ontology, §8.1).

By *its* harnessing as ‘ourselves’ the scientific theory of evolution and the notion of risk (for example), we suppose (in the natural disposition) we are in a better position to understand ourselves. But this supposition, and all the others which constitute us as brain-sign, do not make us *at home* in the world as consciousness supposes. The raw conditions of our being in the world are the risky conditions of our physical being. This is obscured for us by the *unexplained* condition (for example) that we seem to see the world. This is not an illusion the brain creates: it is a condition of its biological function. The brain does not ‘question’ its own creation, except in situations that force it to (in the case of unproductive error, for example). But our supposition that we *do* see the world, and are secure in this sight, is revealed by the scientific disposition as a *circumscription* of our being in the world, biologically determined. This is not a negative statement on our state, but a correction of what the theory of consciousness misleads us into supposing by not being the scientific disposition.

The gravity of the theory of evolution for us is, for the most part, falsely ameliorated by the theory of consciousness, evidenced with such ‘materialist’ theoreticians as Richard Dawkins (§4.4), Daniel Dennett (§3.2.1.3) and Antonio Damasio (§6.3). Each of these writers endorses the theory of evolution. For them it is as permanent a feature as that very universe we see ‘out there’. Though nature is red in tooth and claw, for them *still* we humans rise above it by our sensibilities and sensitivities. Sensibilities and sensitivities deflect the brute determinism of nature.

However, the permanence of the universe we see is a neural expression of the hold the brain has on the universe as its physical states. That *hold* belongs to the physically transient brain. Thus the universe we see as a permanence (a kind of metaphysical supposition of brain-sign itself) is entirely undermined by the transient condition of its speculative creation *as us*. Even more so such theories as evolution.

There is no claim here that the world does not actually exist, or that the theory of evolution is wrong. The point is we cannot rely upon the reality brain-sign seems to endorse *via* our *sense* of being in the world with our sight and our theories.

\*

This might seem a very subtle point, yet it lies at the heart of the difference between a purely neural brain and the error of consciousness. For a hundred years, particularly in the Continental tradition, there has been a withdrawal from the kind of foundationalism that consciousness is supposed to offer (which was the error of Positivism). Brain-sign theory provides a scientific framework for that withdrawal by demonstrating how we function as physical organisms without knowledge. But of course, our relation to ourselves consequently will completely change.

The myriad questions we ask about human nature or personality are not grounded in any viable scientific modus.

How could murderous Nazis love great music?

How could there be child abuse when children are so innocent?

Why, when we are programmed to reproduce, is there homosexuality?

Why can deep love simply evaporate?

Or, for example, to quote Zygmunt Bauman:

‘The unspoken terror permeating our collective memory of the Holocaust...is the gnawing suspicion that the Holocaust could be more than an aberration, more than a deviation from an otherwise straight path of progress, more than a cancerous growth on the otherwise healthy body of the civilised society; that, in short, the Holocaust was not the antithesis of modern civilization and everything (or so we like to think) it stands for. We suspect (even if we refuse to admit it) that the Holocaust could have uncovered another face of the same modern society whose other, so familiar, face we so admire. And that the two faces are perfectly comfortably attached to the same body’ (1979; 7).

Such questionings are scientifically futile for no human nature, personality or civilization exists for which there could be answers. The causes that lie behind such questionings (insofar as we can ask genuine questions) are in the brain *qua* physical states, and only a brain science can reveal them, under the normal caveats of brain-sign theory. That does not stop us asking the questions because we are so built. However, to be more adapted we need to introduce into our questioning the scientific assessment of our biological status.

The brain is an alien because it rules us in ways we cannot devise from the state it creates us as. The character of our lives is determined in a way over which we (as brain-sign) have no control. The point of our lives does not lie in any terms we can conceive – because we do *not* conceive, intellectually.

And so we are at a radical beginning.

# **The Theory of Brain-Sign**

## **10.0 Conclusion**

This thesis seeks to replace consciousness as a theory with a physical theory of the brain phenomenon (brain-sign), involving an explanation of why there was the supposition and theory of consciousness in the first place. Also the impact of brain-sign theory upon theorising generally has been outlined, and the nature of being human. First we summarise briefly, then look to questions and tasks that remain to be explored.

### **10.1 Summary**

Brain-sign theory takes the physical world as what is. Therefore it begins from the physical world without assuming historical suppositions concerning the human relation to God (knowledge, e.g. §3.1.1) or the universe at large (mentalism, e.g. 3.1.2). Moreover the theory seeks a genuine biological rationale for the brain phenomenon expressible in purely physical terms, which it finds already extant in physical biology (signs).

A crucial feature of the theory is the relation of the individual to the operative collective (the super-organism). Brain-sign theory divides the operation of the brain into two (our dual nature), both evidently physical: causality and communication. This clarifies functionally the operation of the brain. Both modalities (causality and communication) are necessary for effective neural operation where dynamic collective action is involved, i.e. that which is uncertain or imprecise. They function in the (inter-organism) physical structure we term the unit of the communicated.

Most particularly, we have introduced a new biological category, brain-sign itself. Although this is analogous to signs elsewhere in the animal kingdom, designed for communal effect, its role is an entirely new biological account of inter-organism operation, signifying the cause of (neural) causality. It underpins the superstructure of brain-sign theory. Additionally it allows a seamless continuity between ourselves and other creatures, reducing our supposed uniqueness in possessing language, for the brain-sign account of language does not accept the mental processes of thought.

We have sketched the three major categories of brain-sign as communication: categories-of-the-world, categories-of-interaction and language. In so doing we have clarified and demarcated the content of brain-sign (as communicative biology) in a way impossible for mentalism. Since the account begins and remains in the physical world, we have been able to eliminate the (physically) irreducible nature of mental states, be they (e.g.) perception or thought, belief or desire, memory or imagination.

The key functional differentiator of brain-sign from consciousness is the removal of knowledge. As shown, knowledge is a mythical instantiation without genuine explanation. Yet it is taken for granted in everyday conduct and (to quote a few disciplines) in logical, scientific and legal activity. But our ability as organism to proceed from situation A to B depends, not upon knowledge, but upon the brain's physical operation (including brain-sign). Thus while it may be the case that the brain can 'do' arithmetic, indeed very complex mathematics, it does not do so by knowledge as consciousness. The signifiatory function of brain-sign is of an entirely different ontological nature from knowledge. A firm grasp of this makes brain-sign theory comprehensible (though it is not a grasp, which would imply knowledge).

In general, we do not act rationally because of the light of consciousness, and as Freud and many before and since have pointed out, our given reasons often do not correspond to what we do, and it can be impossible to determine our reasons for doing things.<sup>137</sup> This is because causation is the brain, and its explanations *as* reasons (i.e. brain-signs), do not possess Reason itself. The philosophical devotion to Reason *qua* consciousness (that which is to a subject) is misplaced. Perhaps, for example, the mathematical status of infinities, the ambiguity in the behaviour of so-called light (as wave or particle) and the Heisenberg Uncertainty Principle demonstrate there are limits to what the brain can do, and that is because the brain does not (as a rational surrogate for the divine status) know anything: its function is for the biological survival of the organism, not to expose the mystical depths of the universe.

A most disquieting aspect is that the process of reading a text (indeed, this text) involves 'appreciating' that none of the arguments has an impact upon the inner mind we are supposed to have. Our causality in the physical universe is of an unreachable neural kind, and

---

<sup>137</sup> Including the seminal study by Nisbett and Ross (1980). See also Patricia Smith Churchland *vis-à-vis* rules of thought/action (2009).

arguments ‘appear’ purely for (neural) inter-organism communication. But this is to ‘appreciate’ both the fact of the physical universe (which the notion of consciousness obscures), and our status as brain-sign. These ‘facts’ remain brain-sign states, and offer us no certainty about, or habitation of the universe as consciousness presupposes. But there is no option in our operation as given.

## 10.2 Future work

Having outlined these topics, we should note briefly the kind of work remaining which, with the groundwork established, is almost everything.

The initial task, though apparently mandatory, is impossible. But in identifying this, subsequent work can proceed, for the context is established. In principle we would like to build the defining relation between neural states and action in the world as scientifically validated knowledge. But this is impossible because there can be no scientifically validated knowledge; there is no knowledge. Although relations between neural states and action might be determined, they must always remain subject to our condition as causal brain and brain-sign. Thus what defines the relation between neural states and action in the world is *our* actual causal neural states which themselves can never aspire to knowledge. It is worth comparing Max Planck here, as quoted by Paul Feyerabend.

‘The two statements, “There exists a real external world which is independent of us” and “This world cannot be known immediately” together form the basis of all physics. However, they are in conflict to a certain extent and thereby reveal the irrational element inherent in physics and in every other science, which is responsible for the fact that a science can never solve its task completely’ (1999; 62).

Max Planck’s stance is a functional approximation to brain-sign theory; but brain-sign theory provides a more comprehensive and coherent account because it does not assume the ‘irrational’ notion of knowledge, with its distinction of internal (mind) and external (world).

Nevertheless, we can only proceed according to our ontological state within the world. Various different streams of investigation lie before us to be charted in different ways. As with any science, we must proceed both theoretically and empirically, anticipating some convergence (cf. §6.12.7.).

1) *Neural to world causality*. What is now evident from the thesis is that the search for mental states in the brain is pointless since they do not exist. Thus the implicit but unestablishable causal powers of mental states cannot be assumed in the brain. But that gives us a new investigative scenario. We have to develop criteria and structure that map the causal capacities of the brain to action in the world (at least as a potentiality). This, which will happen without being knowledge, requires both the establishment of the causal criteria of action, i.e. what the brain sets out to accomplish in causal terms as a theoretical domain, and the physical structure and processes in the brain that correspond to the criteria. Clearly both these are enormous tasks, likely involving new research approaches. Effectively what is required is the delineation of our biological foundations in terms of cause to implementation (perhaps what Spinoza, for example, supposed as *conatus*, though without the notion of mind).<sup>138</sup>

Here is but one issue. Most actions in the world denoting significant ‘intelligence’ by the brain will be hierarchical aggregates of many factors, and these factors, at various levels of the hierarchy, are also employed in parallel or quite other kinds of actions (cf. Bechtel, 2009). Locating and codifying the aggregated segments, and the method of aggregation, obviously involves both a theoretical command of neural activity in its accumulative architecture, and practical methods of penetration to the physical processes of the brain.<sup>139</sup> A problematic aspect is that brains continually extend their previous states in causality to improve (in principle) their self-sustaining. Thus (in time terms) between stone age man and us, and those in another thousand years, brains develop beyond prior foundations, or indeed the foundations themselves develop. Such is the complexity and uncertainty of the adaptational nature of the brain.

Moreover it is well recognised that even what might seem simple tasks, e.g. the tracking of a ball coming to a football player and their kicking it, requires the brain to perform numerous ‘calculations’ both to work out trajectory, and effect the kicking action (not to mention the conditions that place the player in a game in the first place). But the brain

---

<sup>138</sup> The term *conatus* has a long and subtly changing history. As Spinoza characterised it (1910: 91): ‘Everything in so far as it is in itself endeavours to persist in its own being.’ 3PVI, and ‘The endeavour wherewith a thing endeavours to persist in its being is nothing other than the actual essence of the thing.’ 3PVII

In its modern characterisation, biologically, it has been termed *autopoiesis*. E.g. Varela *et al.*, 1974

<sup>139</sup> This is complicated by another factor. As Ira Black states it: ‘The rules of function are the rules of architecture, and function governs architecture, which governs function. We require a new vocabulary, a new set of concepts to replace the functionalist view’ (1991; 8-9). Although Black has stated this, his own text remains dominated by psychologism.

is not functioning as a computer working on partial differential equations. Computers are programmed by human beings with the supposition they calculate understandably in the domain of conscious processes, and can then effect a resulting action based upon the calculation. But the conscious processes by which humans suppose they understand PDEs are actually brain-sign states which have no causal power. Indeed, it is human *brains* that make computers,<sup>140</sup> not conscious humans. The supposition that computers are intelligible results from the false assumption that we know how they work, which in turn results from the false assumption that we know anything. Clearly brains make ‘intelligent’ machines, but brains do not know how they do it: their doing is entirely embedded in their neural processes and the actions they cause. As such (as a generality) brains do not perform a calculation from which they then initiate action. The ‘calculation’ is an upstream part of action causation. Though evidently they can brain-sign calculation as a communicable status (but not the method of calculation upon which the brain itself acts), this must be seen as a necessity of communication (for the super-organism), not action itself. Thus so-called knowledge of computers and their mode of action results from the human as part of a super-organism (as which computers are made employing communicable brain-signs), not as knowledge-based individuals. In terms of the individual brain, action is a seamless process. Thus...

2) *Brain-sign*. Another research programme is the activity of both forming a comprehensive account of what brain-signs signify, i.e. what their content actually embodies or connotes, and how and where in the brain this happens. The first is initially a theoretical programme, the second empirical.

We have commenced a structural account of brain-sign itself, but the topic is enormous. We must suppose the content of brain-sign reflects the essence of the brain’s communicative capability concerning its causal provocation (as established in evolution), but this is clearly not given wholly at one time, nor indeed could its description be finished (as indicated above concerning causality). Even so, further development would undoubtedly add to the reconstruction of our grasp of the human being, and other organisms, in their evolutionary difference and progression (although it is not a grasp in fact).

---

<sup>140</sup> With of course the body and world technical resources.

It is evident that brain-sign types, resulting from causal actuality, repetition and reformulation, although linked in mechanism, are to be differentiated in the nature of the brain's causal currency.<sup>141</sup>

But we must continually take into account that the theory being developed (by us) does not result from the ability of our minds to see and contemplate the material of brain-sign (*contra* e.g. Block on consciousness). We are brain-sign, and in developing the theory, our brains are us (as the theory) in their communicative manifestation. The danger philosophers have seen of our being misled by our imagination *vs.* our grasp of reality (e.g. by Descartes, Spinoza and Kant) is already to have supposed a clean differentiation can be made between imagination and reality. What must regulate our theory development actually is twofold: Firstly we must hold to the notion of brain-sign as a physical sign, whereas the action of our brains in theory development is unreachable by us as brain-sign; and secondly, as a scientific process, our findings must be brought into correspondence with empirical work. Concerning the latter, we have noted in the text that neuroscientists (and others) simply misread what they suppose they find empirically because they presuppose an unsubstantiated model, *viz.* mentalism. Brain-sign theory sets us off on an entirely new path, the regulation of which is clearly identified in advance by the theory, even though the task is intrinsically difficult.

One aspect has been signposted for empirical research. Because we are not conscious there is no question of trying to find where in the brain all the content of brain-sign comes together (Kant's *transcendental unity of apperception*). We have referred to brain-sign's physical status as the virtual array, but from the outset there has been no proposal that this is an actual array viewed by a subject (hence virtual). Being brain-sign (e.g. 'appearing' to see the world) gives us no clue as to the physical location(s) or mechanics as which we are made. Neither chameleon nor stick insect have any notion of their signficatory camouflage, and while the camouflages are (to us) locatable and identifiable as mechanisms *per se*, in the case of brain-sign we are forced to confront the fact that we do not *know* what the camouflage of the chameleon or stick insect are or how they work, so even less could we know about brain-sign. However this will not, and should not deter the development of theories of brain-sign by our brains. We must 'rely' on the physical brain to work out its own (brain-sign) location and mechanics.

---

<sup>141</sup> We noted that Kant 'shrank back' from identifying what was at issue here, §2.10.

3) *Language*. A key topic for future research is language. Clearly brain-sign theory points to a new approach since we will be looking, on the one hand, for language as the impact of external (language) action on the brain as well as action by the organism from the brain on another (i.e. as structured compression waves), and on the other hand as a component of brain-sign itself. Moreover a fascinating aspect is that language operates in different modalities, through ears and the eyes and, for the blind, fingers. Since we have removed the notions of seeing and hearing (and touch), and meaning and understanding, work will concentrate on the physical processes of language-impact as causal. Whatever structure the brain-sign version of language has – which is different between the numerous extant languages – it must result from the brain’s interpretation of those physical processes.<sup>142</sup> It is clear, too, that there is a linkage between what words signify in relation to the other components of brain-sign (categories-of-the-world and -interaction), although they do not result from them, for they all derive from the interpreted causal source. Indeed, to identify language as action rather than an hermetically sealed mental function (with unspecified relation to other functions) is to alter fundamentally the research approach.

Moreover, the ‘verbal’ productions of other animals assumes new significance. For example, at §6.13 we stated that ‘the sound of [animal] emotional calls, which Pinker refers to, are undoubtedly categories-of-the-world precisely because they emanate from the animals in the world creating them in bio-functional terms.’ Yet at §7.5.1 we said: ‘Words appear to be objects “out there”. In brain-sign terms they might seem to be categories-of-the-world because, for example, there they are on the page like a table or chair in the room. But this completely fails to grasp the function of words (and language) as a brain-sign category. Of course the marks on the page are a category-of-the-world. But before we “see” or “hear” *words*, we are already causally orientated toward the world by them in a way completely invisible to us, and this accounts for the fact that we *appear* to know what a word means, and that we *appear* to understand how to use it.’

It is precisely the question of how animals function that is the differentiator here. It seems (to us) highly likely that apes react to communication with a sense of understanding because they are so close to us as primates. But what of whales and dolphins which have

---

<sup>142</sup> Cf. Nicholas Evans and Stephen Levinson (2009; 429). ‘Talk of linguistic universals has given cognitive scientists the impression that languages are all built to a common pattern. In fact, there are vanishingly few universals of language in the direct sense that all languages exhibit them. Instead, diversity can be found at almost every level of linguistic organization. This fundamentally changes the object of enquiry from a cognitive science perspective.’

highly elaborate communication mechanisms, and even in bird song? Does the brain-sign content of these species (assuming birds have brain-sign function) include a sense of understanding – i.e. as a category-of-interaction so projected as brain-sign? Clearly bees have a language in that they can alter each other's behaviour by their dances; but there is no possibility their functioning includes a sense of understanding, for they do not have brain-sign. Theirs is static not dynamic communication.

## The Theory of Brain-Sign

### Bibliography

- Allen, R. and Reber, A.S. (1998) 'Unconscious Intelligence' entry, in *A Companion to Cognitive Science*, W. Bechtel and G. Graham (eds.), Blackwell, 314-323
- Aristotle (1995) *De Anima, The Complete Works of Aristotle*, J. Barnes (ed.), vol. 1, trans. J.A. Smith, Princeton University Press
- Arbib, M. (2005) 'The Mirror System Hypothesis', in Tellerman (2005), 21-47
- Arnauld (1990) *On True and False Ideas*, trans. S. Gaukroger, Manchester University Press
- Baars, B. (1996a) 'Understanding Subjectivity: Global Workspace Theory and the Resurrection of the Observing Self', *Journal of Consciousness Studies*, 3, 211- 216
- Baars, B. (1996b) *In the Theater of Consciousness: The Workspace of the Mind*, Oxford University Press
- Baars, B. (2004). 'A Stew of Confusion', *Journal of Consciousness Studies* (2004) vol. 11, no. 1, 29-31
- Bachelard, G. (1971) *On Poetic Imagination and Reverie*, trans. C. Gaudin, Spring Publications
- Bauman, Z. (1979) *Modernity and the Holocaust*, Polity Press
- Bear, M.F., Connors, B.W., Paradiso, M.A. (2006) *Neuroscience: Exploring the Brain*, 3<sup>rd</sup> edition, Lippincott, Williams and Wilkins
- Beaumont, J.G., Kenealy, P.M., Rogers, M.J.C. (1999) *Dictionary of Neuropsychology*, 2<sup>nd</sup> edition, Blackwell, p527
- Bechtel, W. (2008) *Mental mechanisms: Philosophical perspectives on the sciences of cognition and the brain*, Routledge
- Bechtel, W. (2009) 'Molecules, Systems, and Behaviour: Another View of Memory Consolidation', in Bickle 2009, 13-40
- Bechtel, W., Graham, G., ed. (1998) *A Companion to Cognitive Science*, Blackwell
- Bechtel, W. & Abrahamsen, A. (2002) *Connectionism and the Mind*, 2<sup>nd</sup> edition, Blackwell
- Bennett, M., Dennett, D., Hacker, P., Searle, J., Robinson, D. (2007) *Neuroscience & Philosophy*, Columbia University Press

- Bennett, M., Hacker P. (2003) *The Philosophical Foundations of Neuroscience*, Blackwell
- Benfey, O.T. (1958) 'August Kekulé and the Birth of the Structural Theory of Organic Chemistry in 1858', *Journal of Chemical Education*, 35, 21-23
- Bergson, H. (1911) *Creative Evolution*, trans. A. Mitchell, Henry Holt and Company
- Bergson, H. (1991) *Matter and Memory*, trans N.M. Paul, W.S. Palmer, Zone Books
- Bermúdez, J.L. (2003) 'The Elusiveness Thesis, Immunity to Error through Misidentification, and Privileged Action', in *Privileged Action*, B. Gertler (ed.), Ashgate, 213-231
- Bickle, J. (2006) 'Reducing mind to molecular path ways: explicating the reductionism implicit in current cellular and molecular neuroscience', *Synthese*, 151:411-434
- Bickle, J. (2009) (Ed.) *The Oxford Handbook of Philosophy and Neuroscience*. Oxford University Press
- Black, I.B. (1991) *Information in the Brain: A Molecular Perspective*, MIT Press
- Block, N. (1986) 'Advertisement for a semantics for psychology', *Midwest studies in philosophy*, X; reprinted in: *Mental Representation: A reader* (1994) S. Stich & T. Warfield (eds.), Blackwell, pp 81-141
- Block, N. (1994) 'Consciousness' entry, *A Companion to the Philosophy of Mind*, Blackwell, 210-219
- Block, N. (1995) 'On a confusion about a function of consciousness' *Behavioral and Brain Sciences* 18(2), 227-287
- Block, N. (2001) 'Paradox and cross purposes in recent work on consciousness', *Cognition* 79, 1&2, 197-219; volume published as (2001) *The Cognitive Neuroscience of Consciousness*, N. Dehaene ed., MIT Press
- Block, N. (2007) 'Consciousness, accessibility, and the mesh between psychology and neuroscience', *Behavioral and Brain Sciences*, 30, 5/6, 481-548
- Brannon, E.M., Terrace, H.S. (1998) 'Ordering of the Numerosities 1 to 9 by Monkeys', *Science*, 282, 746-749
- Brecht, B. (1980) *Life of Galileo*, trans. John Willett, Methuen
- Brentano, F. (1973) *Psychology from the Empirical Standpoint*, trans. A. Rancurello, D. Terrell, L. McAllister, Routledge and Kegan Paul
- Byrne, R. M. J. (2005) *The Rational Imagination: How People Create Alternatives to Reality*, MIT Press

- Byrne, R.M.J. (2007) 'Précis of *The Rational Imagination: How People Create Alternatives to Reality*', *Behavioral and Brain Sciences*, 30, 439-480
- The Cambridge History of Later Medieval Philosophy* (1982) N. Kretzmann, A. Kenny and J. Pinborg (eds.), Cambridge University Press
- Burge, T. (2007) 'Psychology supports independence of phenomenal consciousness', *Behavioral and Brain Sciences*, 30, 5/6, 481-548
- Cadava, E., Connor, P., Nancy, J.-L. (1991) *Who Comes After the Subject?*, Routledge
- Caldwell, R., [www.berkeley.edu/aquarius/](http://www.berkeley.edu/aquarius/)
- Carnap, R. (1956) 'The methodological character of theoretical concepts', *Minnesota Studies in the Philosophy of Science*, vol. 1: *The Foundations of Science and the Concepts of Psychology and Psychoanalysis*, ed. H. Feigl and M. Scriven, 38-76
- Chalmers, D. (1995-7) 'Facing up to the Problem of Consciousness', in *Explaining Consciousness: The Hard Problem* (1995-7) Jonathan Shear (ed.), A Bradford Book, MIT Press, and the Journal of Consciousness Studies
- Chalmers, D. (1996) *The Conscious Mind*, Oxford University Press
- Cartwright, N. (1999) *The Dappled World: A Study of the Boundaries of Science*, Cambridge University Press
- Chomsky, N. (1959) Review of Skinner's *Verbal Behavior*, *Language*, 35, 26-58
- Chomsky, N. (2000) *New Horizons in the Study of Language and the Mind*, Cambridge University Press
- Chomsky, N. (2006) *Language and Mind*, 3<sup>rd</sup> edition, Cambridge University Press
- Chambers 20<sup>th</sup> Century Dictionary* (1983) Kirkpatrick, E.M. (ed.), 1983 edition
- Churchland P.M. (1981) 'Eliminative Materialism and the Propositional Attitudes', *The Journal of Philosophy*, 78, 67-90
- Churchland, P.M. (2007) *Neurophilosophy at Work*, Cambridge University Press
- Churchland, P.S. (2009) Inference to the Best Decision. In Bickle 2009, 419-430
- Clark, A. (1996) 'Three varieties of visual field', *Philosophical Psychology*, 9, 477- 495
- Clark, A. (2000) *A Theory of Sentience*, Oxford University Press
- Clark, A. (2001a) 'Visual Experience and Motor Action: Are The Bonds Too Tight', *The Philosophical Review*, 495-519
- Clark, A. (2001b) *Mindware: An introduction to the philosophy of cognitive science*, Oxford University Press

- Clark, A (2002) 'Is Seeing All It Seems? Action, Reason and the Grand Illusion', *Journal of Consciousness Studies*, vol. 9, No. 5/6, 181-201
- Claxton, G. (1999), 'Whodunnit? Unpicking the "Seems" of Freewill', *Academic Press* (99-113)
- Cohen, N. and Squire, L. (1980) 'Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that', *Science* 210, 207-209
- Comrie, B., Kuteva, T. (2005) 'The evolution of grammatical structures' in Tallerman (2005), 185-207
- Crease, R. (2008) *The Great Equations: Breakthroughs in Science from Pythagoras to Heisenberg*. W. W. Norton & Co.
- Crick, F., Koch, C., (1990) 'Towards a neurobiological theory of consciousness', *Seminars in Neuroscience* 2: 263-275
- Crick, F. (1994) *The Astonishing Hypothesis: The Scientific Search for the Soul*, New York: Charles Scribner's Sons
- Cronin, T.W., Marshall, N.J. (1989) 'A retina with at least ten spectral types of photoreceptors in a mantis shrimp', *Nature*, 137-140
- Cummins, R. (1989) *Meaning and Mental Representation*, MIT Press.
- Cummins, R. (1996) *Representations, Targets and Attitudes*, MIT Press
- Damasio, A. (1999) *The Feeling of what Happens*, William Heineman
- Davidson, D. (1975) 'Thought and Talk', in: S. Guttenplan (ed.), *Mind and Language*, Clarendon Press
- Davies, M., Stone, T. (2005) 'Simulation Theory', in *The Shorter Routledge Encyclopedia of Philosophy*, 961
- Dawkins, R. (1989) *The Selfish Gene*, Oxford University Press
- Dawkins, R. (2006) *The God Delusion*, Boston: Houghton Mifflin
- Dennett, D.C. (1969) *Content and Consciousness*, Routledge and Kegan Paul
- Dennett, D.C. (1991) *Consciousness Explained*, Little, Brown & Company
- Dennett, D.C. (1994) 'Dennett' entry, in *A Companion to the Philosophy of Mind*, S. Guttenplan (ed.), Blackwell, 236-244
- Dennett, D.C. (1996) *Kinds of Minds: Towards an Understanding of Consciousness*, Weidenfeld and Nicolson

- Descartes, R. (1968) *Meditations*, trans. F.E. Sutcliffe, Penguin
- Descartes, R. (1985) *The Philosophical Writings of Descartes*, trans. J. Cottingham, R. Stoothoff, D. Murdoch, vol. 1, Cambridge University Press
- Di Pellegrino, J., Klatzky, R., McCloskey, B. (1992) 'Time course of preshaping for functional responses to objects', *Journal of Motor Behavior*, 21, 307-316
- Dilthey, W. (1954) *The Essence of Philosophy*, trans. S.A. Emery & W.T. Emery, University of North Carolina Press
- Dretske, F. (2003) 'How Do You Know You are Not a Zombie?', in *Privileged Access*, ed. B. Gertler, Ashgate, 1-14
- Edelman, G. & Tononi, G. (2000) *Consciousness: How Matter Becomes Imagination*, Penguin
- Edwards, J.C.W. (2005) 'Is Consciousness Only the Property of Individual Cells?' *Journal of Consciousness Studies*, 4-5, 60-76
- Eliot, T.S. (1963) *Ash Wednesday*, from *The Collected Poems 1909-1962*, Faber & Faber Ltd.
- Evans, N., Levinson, S. (2009) 'The Myth of Language Universals: Language diversity and its importance for cognitive science', *Behavioral and Brain Sciences*, 32(5), 429-492
- Feyerabend, P. (1999) *Conquest of Abundance: A Tale of Abstraction versus the Richness of Being*, University of Chicago Press
- Fichte, J.G. (1970) *The Science of Knowledge*, trans. P. Heath & J. Lachs, Meredith Corporation, (1982) Cambridge University Press
- Fields, H.L., Price, D.D. (1994) 'Pain' entry in: *A Companion to the Philosophy of Mind*, Blackwell, 452-459
- Fodor, J. (1975) *The Language of Thought*, Thomas Y. Crowell Co.
- Fodor, J.A. (1980) 'Methodological Solipsism', *Behavioral and Brain Sciences*, 3, 63-109
- Fodor, J.A. (1987) *Psychosemantics*, MIT Press
- Fodor, J.A. (1994) *The elm and the expert*, MIT Press
- Fodor, J.A. (2009) 'It ain't in the head: Jerry Fodor asks if Externalism has outlived its usefulness', *Times Literary Supplement*, [entertainment.timesonline.co.uk](http://entertainment.timesonline.co.uk)
- Frege, G. (1970) 'On Sense and Reference', in *Translations from the Philosophical Writings of Gottlob Frege*, trans. P. Geach, M. Black, Basil Blackwell, 56-78

- Freud, S. (1969) *Introduction to Psycho-Analysis*, trans. J Strachey, The International Psycho-Analytic Library, No. 35, The Hogarth Press and The Institute of Psycho-Analysis
- Freud, S. (1950) *Project for a Scientific Psychology*, The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume I (1886-1899): Pre-Psycho-Analytic Publications and Unpublished Drafts, 281-391
- Freud, S. (1954) *The Origins of Psycho-Analysis, Letters to William Fliess, Drafts and Notes: 1887-1902*, M. Bonaparte, A. Freud and E. Kris (eds.), trans. E. Mosbacher and J. Strachey, Basic Books
- Freud, S. (1960) *The Psychopathology of Everyday Life*, trans. A. Tyson, J. Strachey (ed.), Ernest Benn Ltd.
- Freud, S. (1962a) *The Ego and the Id*, trans. J. Riviere, revised J. Strachey, International Psychoanalytic Library, No. 12, The Hogarth Press and The Institute of Psychoanalysis
- Freud, S. (1962b) *The Future of an Illusion*, trans. W.D. Robson-Scott, revised J. Strachey, The Hogarth Press and The Institute of Psycho-Analysis
- Freud, S. (1963, 1964) *The Complete Introductory Lectures on Psychoanalysis*, trans. J. Strachey, George Allen and Unwin
- Frisch, K. Von (1967) *The dance language and orientation of bees*, Belknap Press of Harvard University Press
- Frith, C. (2007) *Making up the Mind: How the Brain Creates our Mental World*, Blackwell
- Gazzaniga, M.S. (1999) *The Mind's Past*, University of California Press
- Gazzaniga, M.S. (2005) *The Ethical Brain*, Dana Press
- Gaukroger, S. (1990) 'Introduction' to Arnaud's *On True and False Ideas* (1990) Manchester University Press
- Gertler, B. (2006) 'Consciousness and Qualia Cannot be Reduced', *Contemporary Debates in Cognitive Science*, R.J. Stainton (ed.), Blackwell Publishing, 202-216
- Gibson, J.J. (1977). 'The theory of affordances', in R. Shaw & J. Bransford (eds.), *Perceiving, Acting and Knowing*. Hillsdale, NJ: Erlbaum
- Giere, R.N. (2000) 'Theories', in *A Companion to the Philosophy of Science*, W.H. Newton-Smith (ed.), Blackwell, 515-524
- Giere, R. (2006) *Scientific Perspectivism*, University of Chicago Press
- Graham-Rowe, D. (1998) 'Think and it's done', *New Scientist*, 17 Oct 1998

- Greene, B.R. (2004) *The Fabric of the Cosmos: Space, Time and the Texture of Reality*, Allen Lane: Penguin
- Groddeck, G. (1961) *The book of the It*, Mentor Books, The New American Library of World Literature
- Guttenplan, S. (1994) 'Introduction', *A Companion to the Philosophy of Mind*
- Haack, S. (2006) *Pragmatism Old and New: Selected Writings*, Prometheus Books
- Hardin, C.L. (1993) *Color for Philosophers: Unweaving the rainbow*, Hackett
- Harlow, J.M. (1848) 'Passage of iron rod though the head', *Boston Medical and Surgical Journal*, 39, 389-393
- Hatfield, G. (2009) *Perception and Cognition: essays in the philosophy of psychology*, Oxford
- Hauser, M. (1997) *The Evolution of Communication*, A Bradford Book, MIT Press
- Hauser, M. (2009) 'The Mind', *Scientific American*, September 44-51
- Hayman, R. (1986) *Writing Against: A Biography of Sartre*, Weidenfeld and Nicholson
- Heidegger, M. (1962) *Being and Time*, trans. J. Macquarrie and E. Robinson, Basil Blackwell
- Heidegger, M. (1969) 'The End of Philosophy and the Task of Thinking', from *On Time and Being*, trans. J. Stambaugh, Harper and Row (1972) 55-73. In *Martin Heidegger: Basic Writings* (1978) Routledge & Kegan Paul
- Heidegger, M. (1975) *Early Greek Thinking*, trans. D.F. Krell & F.A. Capuzzi, Harper and Row
- Heidegger, M. (1982) *The Basic Problems of Phenomenology*, trans. A. Hofstadter, Indiana University Press
- Heidegger, M. (1990) *Kant and the Problem of Metaphysics*, trans. R. Taft, Indiana University Press
- Heil, J. (2003) *From an Ontological Point of View* Clarendon Press, Oxford University Press
- Hilbert, D.R. (1992) 'What is color vision?' *Philosophical Studies*, 68, 351-370
- Hofstadter, D. (2007) *I am a Strange Loop*, Basic Books
- Horwich, P. (2003) *Meaning*, extracted in *Meaning* Richard, M. (ed.), Blackwell Publishing, from *Meaning* (1998) Oxford: Clarendon Press.
- Horwich, P. (2005) *Reflections on Meaning*, Oxford: Clarendon Press

- Horwich, P. (2010) *Truth-Meaning-Reality*, Oxford: Clarendon Press
- Hume, D. (1962) *A Treatise of Human Nature: Book One*, D.G.C. McNab (ed.), Collins
- Hurley, S.L. (1998) *Consciousness in Action*, Harvard University Press
- Hurley, S. (2008) ‘The shared circuits model (SCM): How control, mirroring, and simulation can enable imitation, deliberation, and mind reading’, *Behavioral and Brain Sciences*, 31, 1-58
- Husserl, E. (1960) *Cartesian Meditations*, trans. D. Cairns, Kluwer Academic Publications
- Husserl, E. (1982) *Ideas Pertaining to a Pure Phenomenology and to Phenomenological Psychology, book one*, trans. F. Kersten, Kluwer Academic Publishers
- Jackson, F. (1986) ‘What Mary Didn’t Know’, *The Journal of Philosophy*, LXXXII: 291-295
- Jeannerod M. (1997) *The Cognitive Neuroscience of Action*, Blackwell
- Jeannerod, M. (2006) *Motor Cognition: What actions tell the self*, Oxford University Press
- Jones, E. (1964) *The Life and Work of Sigmund Freud*, abridged edition as a Pelican book
- Jordan, S., McBride, D. (2007) ‘Stable Instabilities in the Study of Consciousness’, *Journal of Consciousness Studies* (2007), vol. 14, no. 1-2
- Jung, C.G. (1969) *The Structure and Dynamics of the Psyche*, trans. R.F.C. Hull, second edition, Routledge and Kegan Paul
- Kaminski, J., Call, J., Fisher, J. (2004) ‘Word learning in a domestic dog: Evidence for “fast mapping”’, *Science* 304, 1682-1683
- Kane, R. (2002) *The Oxford Handbook of Free Will* ‘Introduction’ & ed. R. Kane, Oxford University Press
- Kant, I. (1933) *The Critique of Pure Reason*, 2<sup>nd</sup> edition, trans. N.K. Smith, Macmillan Press
- Kant, I. (1971) *Prolegomena to Any Future Metaphysics*, trans. J.W. Ellington, Hackett Publishing Company
- Kelber, A., Osorio, D., Vorobyev, M. (2003) ‘Animal colour vision: behavioural tests and physiological concepts’, *Biological Reviews*, Cambridge Philosophical Society, (78) 1, 81-118
- Kiesel, T. (1993) *The Genesis of Heidegger’s Being and Time*, University of California Press

- Kiesel, T. (2002) 'Why Students of Heidegger Will Have to Read Emil Lask' in *Heidegger's Way of Thought*, A. Denker & M. Heinz (eds.), Continuum, 101-136
- Kim, J. (2005) *Physicalism, or Something Near Enough*, Princeton University Press
- Kirk, G.S., Raven, J.E., Schofield, M. (1983) *The Presocratic Philosophers*, 2<sup>nd</sup> edition, Cambridge University Press
- Knudsen, C. (1982) 'Intentions and Impositions', *The Cambridge History of Later Medieval Philosophy*, N. Kretzmann, A. Kenny and J. Pinborg (eds.)
- Kuhn, T. (1993) 'Afterwords', in *World Changes: Thomas Kuhn and the Nature of Science*, ed. P. Horwich, A Bradford Book, MIT Press, 311-341
- Kuhn, T. (1996) *The Structure of Scientific Revolutions*, 3<sup>rd</sup> edition, University of Chicago Press
- Land, E.H. (1977) 'The Retinex Theory of Color Vision', *Scientific American*, 108-129
- Laplanche, J., Pontalis, J.B. (1973) *The Language of Psychoanalysis*, The Hogarth Press Ltd.
- Lashley, K. (1951) 'The problem of serial order in behavior', *Cerebral mechanisms in behavior*, L.A. Jeffress (ed.), Wiley 112-146
- Lauer, Q. (1965) 'Introduction' to Husserl's *Phenomenology and the Crisis of Philosophy*, trans. Q. Lauer, Harper TorchBooks
- Lawrence, D.H. (1950), *D.H. Lawrence Selected Poems*, Penguin
- Lehrer, J (2011) *Proust was a Neuroscientist*. Cannongate
- Leplin, J. (2000) 'Realism and Instrumentalism', in *A Companion to the Philosophy of Science*, as above, 393-401
- Levine, J. (2003) 'Experience and Representation' in *Consciousness: New Philosophical Perspectives*, Q. Smith and A. Jokic (eds.), Oxford University Press, 57-76
- Levitin, D. (2006) *This Is Your Brain On Music: Understanding a Human Obsession*, Dutton Penguin
- 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, pp 623-642
- Libet, B. (1999) 'Do We Have Free Will?', *Journal of Consciousness Studies*, 6, No. 8-9, 47-57
- Lieberman, P. (2006) *Toward an Evolutionary Biology of Language*, Belknap Harvard

- Locke, J. (1997) *An Essay Concerning Human Understanding*, R. Woolhouse (ed.), Penguin
- Lycan, W. (2006) 'Consciousness and Qualia Can be Reduced', in *Contemporary Debates in Cognitive Science*, R.J. Stainton (ed.), Blackwell Publishing, 189-201
- Lyn, H. (2007) 'Mental representation as revealed by vocabulary errors in two bonobos (*Pan paniscus*)', *Animal Cognition*, 10, 461-475
- Machamer, P. (2009) Learning, Neuroscience, and the Return of Behaviorism, 166-178. In Bickle 2009
- Macphail, E.M. (1998) *The Evolution of Consciousness*, Oxford University Press
- Makkreel, R. (1990) *Imagination and Interpretation in Kant*, Chicago University Press
- Malebranche, N. (1980) *The Search after Truth and Elucidations of the Search after Truth*, trans. T.M. Lennon and P.J. Olscamp, Columbus: Ohio State University Press
- Marr, D. (1983) *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman
- Martin, A., Haxby, J.V., Lalonde, F.M., Wiggs, C.L., Ungerleider, L.G. (1995a) 'Discrete cortical regions associated with knowledge of color and knowledge of action', *Science*, 270:102-105
- Martin, A. Wiggs, C.L., Ungerleider, L.G., Haxby, J.V. (1995b) 'Neural correlates of category-specific knowledge', *Nature*, 379, 649-652
- Matthen, M. (2001) 'What Colors? Whose Colors?', *Consciousness and Cognition*, 10, 117-124
- McCrone, J. *Going Inside* (1999), Faber and Faber
- McDowell, J. (1996) *Mind and World*, Harvard University Press
- McGinn, C. (1989) 'Can We Solve the Mind-Body Problem?' (1989) *Mind*, Volume 98; reprinted in (1995) *Modern Philosophy of Mind* ed. W. Lyons, 272-295
- Mead G.H. (1934) *Mind, Self and Society from the Standpoint of a Social Behaviorist*, C.W. Morris (ed.), Chicago University Press
- Metzinger, T. (2003) *Being No One*, MIT Press
- Metzinger, T. (2005) 'Précis: Being No One', *Psyche*, 11 (5)
- Miller, G.A (1964) *Psychology: The Science of Mental Life*, Hutchison and Co. Ltd.
- Miller, G.A. (2003) 'The cognitive revolution: a historical perspective', *Trends in Cognitive Science*, Vol. 7, No. 3

- Millikan, R.G. (1984) *Language, Thought and Other Biological Categories*, MIT Press
- Milner, D., Goodale, M. (2004) *Sight Unseen*, Oxford University Press
- Milner, A.D., Goodale, M.A. (2006) *The Visual Brain in Action*, 2<sup>nd</sup> edition, Oxford University Press
- Mishkin, M., Ungerleider, (1982) 'Two cortical visual systems', in *Analysis of Visual Behavior*, D. Inge, M. Goodale, R. Mansfield (eds.), MIT Press, 549-586
- Monk, R. (1990) *Ludwig Wittgenstein: The Duty of Genius*, Jonathan Cape
- Nagel, T. (1974) 'What is it like to be a bat?', *Philosophical Review*, 83, 435-50
- Neville, H.J., Bavelier, D., Corina, D., Rauschecker J., Karni, A., Lalwani, A., Braun, A., Clark, V., Jezzard, P., Turner, R. (1998) 'Cerebral organization for language in deaf and hearing subjects: biological constraints and effects of experience', *Proceedings of the National Academy of Sciences USA*, February 3, 95(3), 922-929  
Online: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=33817>
- Nisbett, R.E., Ross, L.E. (1980) *Human Inference: Strategies and shortcomings of social judgement*, Prentice Hall
- Nietzsche, F. (1961) *Thus Spoke Zarathustra*, trans. R.J. Hollingdale, Penguin
- Nietzsche, F. (1986) *Human, All Too Human*, trans R.J. Hollingdale, Cambridge University Press
- Noë, A. (2005) *Action in Perception*, MIT Press
- Norretranders, T. (1991) *The User Illusion: Cutting Consciousness Down to Size*, trans. J. Sydenham, Penguin Books
- Olafson, F. (1987) *Heidegger and the Philosophy of Mind*, Yale University Press
- O'Shaughnessy, B. (1992) 'The Diversity and Unity of Action and Perception', in *The Contents of Experience*, T. Crane (ed.), Cambridge University Press, 216-266
- Paller, K.A., Ranganath, C., Gonsalves, B., LaBar, K.S., Parrish, T.B, Gitelman, D.R., Mesulam, M.M., Reber, P.J. (2003) 'Neural Correlates of Person Recognition', *Learning and Memory*, Cold Spring Harbor Laboratory Press, 10: 253-260
- Papineau, D. (2003) 'Theories of Consciousness' in *Consciousness: New Philosophical Perspectives*, ed. Q. Smith and A. Jokic, Oxford University Press, 353-383
- Parfit, D. (1984) *Reasons and Persons*, Oxford University Press

- Peirce, C.S. (1931-66) *Collected Papers*, vol. 1-8, C. Hartshorne, P. Weiss, A. Burke (eds.), Harvard University Press
- Penfield, W. (1958) *The Excitable Cortex in Man*, Liverpool University Press
- Peschl, M.F. (1999) 'The development of scientific concepts and their embodiment in the representational activities of cognitive systems', *The Nature of Concepts: Evolution, structure and representation*, ed. P. Van Loocke, Routledge, 184- 214
- Pinker S. (1997) *How the Mind Works*, Allen Lane, The Penguin Press
- Plato (1974) *The Republic*, trans. D. Lee, 2<sup>nd</sup> edition, Penguin
- Plato (1975) *Phaedrus & Letters VII and VIII*, trans. W. Hamilton, Penguin
- Prinz, J., Clark, A. (2004) 'Putting Concepts to Work: Some Thoughts for the Twentyfirst Century' (2004) *Mind and Language*, Vol. 19, No.1, 57-69
- Putnam, H. (1975) 'The meaning of "meaning" in: K. Gunderson (ed.), *Language, mind and knowledge: Minnesota studies in the philosophy of science*, Volume 7, Minneapolis: University of Minnesota Press
- Pylyshyn, Z. (2003) *Seeing and Visualising: It's not what you think*, A Bradford Book, MIT Press
- Quine, W.V.O. (1951) 'Two Dogmas of Empiricism', *Philosophical Review*; reprinted in *From a Logical Point of View*, second edition revised (1980), Harvard University Press, 20-46
- Quine, W.V.O. (1960) *Word and Object*, MIT Press
- Rey, G. (1983) 'A reason for doubting the existence of consciousness', in *Consciousness and Self-Regulation*, R. Davidson, G.E. Schwartz, and D. Shapiro (eds.), vol. 3, Plenum Press
- Rakover, S.S, Cahlon, B. (2001) *Face Recognition: Cognitive and computational processes*, John Benjamins Publishing Company
- Rey, G. (2007a) 'Phenomenal Content and the Richness and Determinacy of Color Experience', *Journal of Consciousness Studies*, Vol. 14, 9-10, 112-131
- Rey, G. (2007b) 'Resisting Normativism in Psychology', in *Contemporary Debates in Philosophy of Mind*, B.P McLaughlin and Jonathan Cohen (eds.), Blackwell Publishing, 69-84
- Rizzolati, P., Fadiga, L., Fogassi, J. (1996) 'Premotor cortex and the recognition of motor actions', *Cognitive Brain Research*, 3, 131-141
- Rorty, R. (1980) *Philosophy and the Mirror of Nature*, Basil Blackwell

- Rorty, R. (1992) 'Heidegger, Contingency, and Pragmatism' in (1992) *Heidegger: A Critical Reader*, H.L. Dreyfus & H. Hall (eds.), Basil Blackwell Ltd.
- Rose S.P. (2008) 'In search of the God neuron', *Guardian Review*, 27 December 2008
- Rosenberg, A. (2000) 'Social Science, Philosophy of', in *A Companion to the Philosophy of Science*, 451-460
- Rosenthal, S.B. and Bourgeois, P.L. (1980) *Pragmatism and Phenomenology: A Philosophic Encounter*, B.R. Gruner Amsterdam
- Ross, P.W. (2001) 'The Location Problem for Color Subjectivism', *Consciousness and Cognition*, 10, 42-58
- Roy, J.-M. (1999) 'Saving Intentional Phenomena: Intentionality, Representation, and Symbol', in *Naturalizing Phenomenology*, J. Petitot, F.J. Varela, B. Pachoud, J.-M. Roy (eds.), Stanford University Press, 111-147
- Sacks, O. (1985) *The Man who Mistook his Wife for a Hat*. Picador
- Sacks, O. (2007) *Musicophilia: Tales of music and the brain*, Borzoi Books
- Safranski, R. (1998) *Martin Heidegger: Between Good and Evil*, Harvard University Press
- Sartre, J.-P. (1944) *Huis Clos*, trans. S. Gilbert as *No Exit*, Vintage
- Sartre, J.P. (1957) *Being and Nothingness: An Essay on Phenomenological Ontology*, trans. H. Barnes, Methuen
- Sartre, J.P (1965) *Nausea*, trans. R. Baldick, Penguin
- Saussure, F. de (1970) 'On the Nature of Language', reproduced in Lane, M. (ed.) *Structuralism, A Reader*, Jonathan Cape. From *Course in General Linguistics* (1959) trans. W. Baskin, Philosophical Library, New York
- Schacter, D.L. and Graf, P. (1986) 'Effects of elaborative processing on implicit and explicit memory for new associations', *Journal of Experimental Psychology: Learning, Memory and Cognition*, 12(3), 432-444
- Schelling, F.W.J. von (1978) *System of Transcendental Idealism*, trans. P. Heath, Charlottesville University Press of Virginia
- Sellars, W. (1997) *Empiricism and the Philosophy of Mind*, Harvard University Press; first published in *Minnesota Studies in the Philosophy of Science*, 1956
- Shettleworth, S.J. (1998) *Cognition, Evolution and Behavior*, Oxford University Press
- Schopenhauer, A. (1960) *On the Freedom of the Will*, trans. K. Kolenda, Bobbs- Merrill Co. Inc.

- Schopenhauer, A (1969) *The World as Will and Representation*, vol. 1, trans. E.F.J. Payne, Dover
- Seager, W. (1999) *Theories of Consciousness*, Routledge
- Searle, J. (1994) *The Rediscovery of Mind*, MIT Press
- Selfridge, O. (1959) ‘Pandemonium: A Paradigm for Learning’, *Symposium on the Mechanisation of Thought Processes*, HM Stationery Office
- Shoemaker, S. (1988) ‘On Knowing One’s Own Mind’, J. Tomberlin ed., *Philosophical Perspectives 2*, Ridgeway Publishing Company, 183-209
- Slezak, P. (2001) ‘Representation: Where Philosophy Goes When It Dies’, *Proceedings of 23<sup>rd</sup> Annual Conference of the Cognitive Science Society*, 940-945, Lawrence Erlbaum Publishers
- Slezak, P. (2004) ‘The World Gone Wrong? Images, Illusions, Mistakes and Misrepresentations’, in *Representation In Mind: New Approaches to Mental Representation*, H. Clapin, P. Staines, P. Slezak (eds.), Elsevier, 163-185
- Smart, J.J.C. (1959) ‘Sensations and Brain Processes’, *Philosophical Review*, 68, 141-156
- Solms, M., Turnbull, O. (2002) *The Brain and the Inner World*, Karnac Press
- Spengler, O. (1926) *The Decline of the West*, trans. C.F. Atkinson, Alfred A. Knopf – Borzoi Books
- Spinoza, B. (1910) *Ethics*, trans. A. Boyle, Everyman Library
- Sridharan, D., Levitin, D.J., Chafe, C.H., Berger, J., Menon, V. (2007) ‘Neural dynamics of event segmentation in music: Converging evidence for Dissociable Ventral and Dorsal Networks’, *Neuron*, 55, 3
- Stapp, H. (2006) ‘Quantum Interactive Dualism, II: The Libet and Einstein-Podolsky-Rosen Causal Anomalies’, *Erkenntnis*, Vol. 65-1, 117-142
- Sterelny, K. (2003) *Thought in a Hostile World: The Evolution of Cognition*, Blackwell
- Stich, S. (1978) ‘Autonomous psychology and the belief-desire thesis’, *The Monist*, 61, 4, 573-591
- Suppe, F. (2005) ‘Scientific Theories’, in *The Shorter Routledge Encyclopedia of Philosophy*, Routledge, 1016
- Tallerman, M. (ed.) (2005) *Language Origins: Perspectives on Evolution*, Oxford University Press

- Terza, D.d. (1972) 'On Pirandello's Humorism', from *Veins of Humor*, ed. H. Levin, © President and Fellows of Harvard College, reprinted (1989) in *Modern Critical Views: Luigi Pirandello*, ed. H. Bloom, Chelsea House Publishers
- The Shorter Oxford English Dictionary* (1973) 3<sup>rd</sup> edition, Clarendon Press
- Tolstoy, L. (1942) *War and Peace*, trans. L. and A. Maude, Macmillan
- Tye, M. (1995) *Ten Problems of Consciousness*, MIT Press
- Tye, M. (2000) *Consciousness, Color, and Content*, MIT Press
- Tye, M. (2009) *Consciousness Revisited: Materialism Without Phenomenal Concepts*, MIT Press
- Uwechue, O.A. and Pandya, A.S. (1997) *Human Face Recognition Using Third-Order Synthetic Neural Networks*, The International Series in Engineering and Computer Science, Kluwer Academic Publishers
- Varela, F.J., Maturana, H.R., Uribe, R. (1974). Autopoiesis: the organization of living systems, its characterization and a model. *Biosystems* 5 187–196
- Walker, O.J. (1939) 'August Kekulé and the Benzene Problem', *Annals of Science* 4, 34-46
- Wall, P. (1999) *Pain: The Science of Suffering*, Weidenfield and Nicolson
- Warden, C.J., Warner, L.H. (1928) 'The sensory capacities and intelligence of dogs, with a report on the ability of the noted dog "Fellow" to respond to verbal stimuli', *Quarterly Review of Biology*, 3, 1-28
- Watson, G. (1995) 'Free Will' entry, in *A Companion to Metaphysics*, J. Kim & E. Sosa (eds.), Blackwell, 175-182
- Watson, G. ed. (2003) *Free Will*, 2<sup>nd</sup> edition, Oxford University Press
- Wegner, D.M. (2002) *The Illusion of Conscious Will*, MIT Press
- Wiener, N. (1948) *Cybernetics; or, Control and Communication in the Animal and the Machine*. Wiley
- Weiskrantz, L. (1997) *Consciousness Lost and Found: A Neuropsychological Exploration*, Oxford University Press
- White, L.L. (1960) *The Unconscious Before Freud*, Basic Books
- Wittgenstein, L. (1967) *Philosophical Investigations*, 3<sup>rd</sup> edition, trans. G.E.M. Anscombe, Blackwell
- Yolton, Y. (1984) *Perceptual Acquaintance*, Oxford University Press