Murray Shanahan

Global Access, Embodiment, and the Conscious Subject

Abstract: The objectives of this article are twofold. First, by denying the dualism inherent in attempts to load metaphysical significance on the inner/outer distinction, it defends the view that scientific investigation can approach consciousness in itself, and is not somehow restricted in scope to the outward manifestations of a private and hidden realm. Second, it provisionally endorses the central tenets of global workspace theory, and recommends them as a possible basis for the sort of scientific understanding of consciousness thus legitimised. However, the article goes on to argue that global workspace theory alone does not constitute a fully worked-out objective account of the conscious subject. This requires additional attention to be paid to (at least) the issue of embodiment, and to the possibility of indexicality that arises when an instantiation of the global workspace architecture inhabits a spatially localised body.

1. Introduction

The aim of this article is to strengthen the foundations of ongoing attempts to investigate consciousness empirically, first by undermining the dualistic thinking that seeks to limit the scope of such enquiries, and second by presenting an objective account of the conscious subject. The first aim is pursued with the aid of Wittgenstein's private language remarks. The upshot is to deny metaphysical import to the inner/outer distinction — the distinction between, on the one hand, the subjective and experiential and, on the other hand, the objective and physical — and thereby to reconcile the so-called hard and easy problems of consciousness (Chalmers, 1996).[1] The second challenge is met by clarifying the relationship between access consciousness and the experiencing subject, where the

Correspondence:
Murray Shanahan, Department of Computing, Imperial College London, 180 Queen's Gate, London SW7 2AZ, UK. *Email: m.shanahan@imperial.ac.uk*

---

[1] Note that, under the inner/outer distinction in question (which is discussed in detail in Section 3), the internal workings of the physical brain are squarely in the realm of the 'outer'.

concept of access consciousness is derived from Block (1995) and the capacity for it is taken to be roughly co-extensive with instantiating a global workspace architecture (Baars, 1988; 1997).[2] Specifically, the article advances the view that an experiencing subject arises when access consciousness inhabits a spatially localised body. Since this argument is the article's main claim to originality, a few more words to introduce it are in order.

A remarkable property of humans and other animals is that, while alive, our bodies occupy clearly bounded, contiguous regions of space that do not in general overlap with other objects. When a living animal moves around, it does so as one complete lump of matter, and although its limbs are subject to changes in their configuration, they retain their basic shape and remain attached to the animal's body. Moreover, a living animal's body doesn't spread out or break into separate parts (a severed limb is, quite properly, disowned), and it cannot move in two directions at once. Many things follow from these biological fundamentals, which Metzinger calls the 'singular embodiment constraint'.[3] One consequence is that our sensory apparatus, which we carry around with us, is also spatially localised, and is obliged to take in the world from a particular location, that is to say from a particular point of view.

An animal's confinement to a spatially localised body — and the limited window on the world entailed by the attachment of its sensory apparatus to that body — has important further ramifications for the sort of information processing carried out by the control system that has to maintain the integrity and wellbeing of that body. These ramifications relate in an important way to two conditions that must be fulfilled by any instantiation of the global workspace architecture. First, the messages that are broadcast to its parallel specialist processes must be in a form that is potentially intelligible by all of them. Second, the ensemble of parallel specialist processes must serve a common remit.

The central argument of this article hinges on the claim that confinement to a spatially localised body guarantees these conditions in a special sense. In the context of a spatially localised body, broadcast messages are intelligible to all the parallel specialist processes that receive them insofar as these messages always pertain to the situation from *this* viewpoint, according to a sensory apparatus located *here*. Similarly, the ensemble of parallel specialist processes serve a common remit because they are all directed towards the control of *this* body — the body to which the aforementioned sensory apparatus is attached — and their duty is to sustain it.[4] This dual indexicality — of the *situation for* this body from the *viewpoint of* this body — binds the mechanism for access consciousness to

---

[2] The concept of access consciousness is characterised more fully in the next section.

[3] Metzinger (2003), p. 161. As Metzinger observes, 'this basic fact is frequently overlooked'. But he by no means draws the same conclusions from it as the present article. Indeed, he allows for the possibility of 'distributed beings' with 'widely scattered sensory organs' who nevertheless develop 'noncentered phenomenal states'. According to the argument of this paper, such beings cannot constitute conscious subjects.

[4] A related proposal is put forward by Eilan (1995), for whom the consciousness of conscious perceptions depends on their 'essentially perspectival' spatial content. However, global workspace theory is not within the purview of Eilan's paper. Evans's (1982, pp. 151–70) discussion of 'egocentric space'

the spatially localised organism it inhabits. The attitude we take towards access consciousness thus embodied is the attitude we take towards a subject and, in Nagel's (1974) well-known phraseology, we speak of its being like something to be that subject.[5]

2. Phenomenal Consciousness and the Hard Problem

Block (1995) introduces a distinction between two aspects of consciousness that he terms *phenomenal* consciousness (or P-consciousness) and *access* consciousness (or A-consciousness).

> P-consciousness is experience. P-conscious properties are experiential properties. P-conscious states are experiential states, that is, a state is P-conscious if it has experiential properties. The totality of the experiential properties of a state are 'what it is like' to have it (Block, 1995, p. 230).

As is apparent from the above quote, there is a close connection between Block's conception of phenomenal consciousness and Nagel's frequently quoted assertion that 'no matter how the form [of consciousness] may vary, the fact that an organism has conscious experience *at all* means, basically, that it is like something to *be* that organism' (Nagel, 1974, p. 436).

Nagel goes on to present his well known argument that conscious experience is irreducible to the physical because 'every subjective phenomenon is essentially connected with a single point of view, and it seems inevitable that an objective, physical theory will abandon that point of view' (p. 437). We will attend closely to this argument later in the paper. But for now, the focus of the discussion is Block, who specifically warns against the 'dangerous' tendency to conflate phenomenal consciousness with access consciousness.

> A state is access-conscious (A-conscious) if, in virtue of one's having the state, a representation of its content is (1) ... poised to be used as a premise in reasoning, (2) poised for rational control of action and (3) poised for rational control of speech.[6]

Block's conception of access consciousness assumes a very coarse-grained division of labour within the mind. He distinguishes only three processes — action, reason, and speech. By contrast, the corresponding conception in this article

---

is also resonant. According to Evans, the subject 'must conceive himself to *be* somewhere — at a point in the centre of an egocentric space capable of being enlarged so as to encompass all objects' (p. 168), and therefore 'it is not possible for a single subject to think of two (or more) separate places as "here"' (p. 167). Moreover, he claims, 'an egocentric space can exist only for an animal in which a complex network of connections exists between perceptual input and behavioural output' (p. 154).

[5] The wording here deliberately echoes Wittgenstein (1958, p. 178): 'My attitude towards him is an attitude towards a soul. I am not of the *opinion* that he has a soul.' He goes on to remark that 'the human body is the best picture of the human soul'. Substituting the word 'subject' for 'soul' in these reflections strips them of their religious overtones while leaving the insight bare.

[6] Block (1995), p. 231. A somewhat differently worded definition is presented in the revised version of the paper that appears in (Block, *et al*., 1997). But the original version is retained here as it is perhaps more immediately evocative of the concept in question. Block's conception, like that of Metzinger (2003, p. 31), differs in a number of ways from that put forward by Baars. But these minor variations have no impact on the thesis of the present paper.

allows for a much finer grain of decomposition, wherein a complex form of behaviour such as speech production is the result of many brain processes acting in concert and over time. Moreover, Block places great emphasis on rationality, while the version of global workspace theory advocated here appeals more neutrally to brain processes. Despite these differences, the underlying idea carries over. Access consciousness arises when widespread, multiple brain processes have access to the same information, which is another way of saying that the information in question is broadcast to these processes.[7]

According to Block, theories of access consciousness, such as those of Shallice (1988) and Baars (1988), cannot serve as theories of phenomenal consciousness. To reinforce the claim that access and phenomenal consciousness are distinct, Block presents purported examples both of P-consciousness without A-consciousness and of A-consciousness without P-consciousness.[8] Drawing on similar intuitions, Chalmers (1996) distinguishes what he calls the 'hard problem' of consciousness from the 'easy problem'.

> [A lot of writing] addresses the 'easy' problem of consciousness: How does the brain process environmental stimulation? How does it integrate information? How do we produce reports on internal states? These are important questions, but to answer them is not to solve the hard problem: Why is all this processing accompanied by an experienced inner life?[9]

Chalmers discusses several examples of theories that address the 'easy problem' of consciousness yet fail to address the 'hard problem', including Baars's global workspace theory (Baars, 1988). Chalmers concedes that 'the global workspace framework is ... well suited to explaining consciousness in its whole bundle of psychological senses' (Chalmers, 1996, p. 112), but he rejects the possibility that it also explains phenomenal consciousness.

> One might suppose that according to the theory, the contents of experience are precisely the contents of the workspace. But even if this is so, nothing internal to the theory *explains* why it is that the information in the global workspace is experienced. The best the theory can do is to say that the information is experienced because it is globally accessible. But now the question arises in a different form: Why should global accessibility give rise to conscious experience?[10]

---

[7] This sub-personal concept of access should not be confused with a personal level conception, such as that employed by Marcel (2003) when he argues (very persuasively) that 'one does not have direct or transparent access to one's own consciousness'.

[8] Each of Block's alleged counter-examples to the identity of A-consciousness and P-consciousness is faulty, in this author's view, because it lacks a proper treatment of related working memory issues. The same criticism will in due course be made of another supposed counter-example supplied by Bayne & Chalmers (2003).

[9] Chalmers (1996), pp. xi–xii. There is an extensive literature on Chalmers' 'hard problem', dating from Shear's (1997) edited volume to the present day (Aleksander, 2005, ch. 7). There will be no attempt to survey this literature here.

[10] Chalmers (1996), p. 112. Block's recent writings do not favour the use of the A-consciousness / P-consciousness distinction to bolster Chalmers' sceptical position, as he approvingly discusses the possibility of finding the neural correlates of P-consciousness, while maintaining that these are likely to be different from the neural correlates of A-consciousness (Block, 2005).

One aim of this article is to confront Chalmers's question: 'Why should global accessibility give rise to conscious experience?' The overall argument has two steps. The first step is to undermine the dualistic thinking that leads up to the question. Authors who express scepticism over the very possibility of arriving at an objective theory of phenomenal consciousness are typically guilty of putting their sceptical carts before their dualistic horses. An exception will be made for Nagel's argument from the irreconcilability of the objective and the subjective, which will be given separate treatment in step two of the argument.

With the possibility of an empirically informed account of consciousness partially reinstated, the second step of the argument is to show that a suitably embodied instantiation of the global workspace architecture promises such an account. Moreover, the intimate relationship between the global workspace architecture and the spatially confined body under its control suggests a way to accommodate a demystified conception of the first-personal subject within an objective description.

3. The Inner/Outer Distinction

Chalmers' revitalisation and defence of the argument from the 'logical possibility of zombies' has had widespread influence, especially in the context of the scientific study of consciousness, as he uses it to justify the purported distinction between the so-called easy and hard problems (Chalmers, 1996). The argument from the logical possibility of zombies is a variant of the problem of other minds, whose textbook presentation asks us to consider how we can have knowledge of the inner lives of others given only their outward behaviour. In a similarly sceptical vein, Chalmers asks us to entertain the possibility of his zombie twin, who is 'molecule for molecule identical to me, and identical in all the low-level properties postulated by a completed physics, but [who] lacks conscious experience entirely' (Chalmers, 1996, p. 94). The zombie twin is physically, functionally, and psychologically identical to the real Chalmers, yet lacks phenomenal consciousness.

The thought expressed here is more radically sceptical than the traditional problem of other minds because its scepticism is not confined to the consideration of outward behaviour but encompasses the internal workings of the brain. Hence its pertinence to the scientific study of consciousness, whose limit it apparently sets. The contention of the present section is that, rather than constituting a successful argument for a dualistic metaphysics, the language used to express the sceptical thought is inherently dualistic from the outset.[11] This is not an original reply, as it simply restates the essential insight of Wittgenstein's private language remarks. However, it is uncommon for commentators on Chalmers even to mention Wittgenstein, let alone to engage fully with his ideas,

---

[11] A similar strategy is adopted by Dennett (1997) when he responds to Block's A-consciousness/ P-consciousness distinction in defence of the strategy of Dennett (1991).

and many are open to the charge that they are trapped within the prison of dualism from which a proper reading of Wittgenstein would liberate them.[12]

Let us rehearse the private language remarks (Wittgenstein, 1958) with Chalmers' argument in mind. Wittgenstein sets the scene by asking in what sense our sensations are 'private' (§246). A protagonist such as Chalmers might assert that only he really knows, for example, when he is in pain. Others can at best surmise that he is in pain from his behaviour. So, adopting his protagonist's turn of language and this attendant notion of privacy, Wittgenstein invites us to consider the possibility of a language 'which describes my inner experiences and which only I myself can understand' (§256). It will be a language that only he can understand because it describes his supposed sensations 'in themselves', and has no relation to any public expression or manifestation of those sensations.

To get a better grip on the idea of such a language, Wittgenstein asks us to imagine a diary in which he records all the days on which he has a certain sensation. To do this, he invents a symbol, say 'S', and in his diary he marks each day on which he has that sensation with this symbol. Now, Wittgenstein asks, how would he know whether he was marking the diary correctly? How would he know that he was using the symbol 'S' for the same sensation on different occasions? If the sensation is truly 'private', what criterion for *correctness* could there possibly be? Marking the diary, Wittgenstein contends, is nothing but an empty ceremony: 'One would like to say: whatever is going to seem right to me is right. And that only means that here we can't talk about right' (§258).

At first sight, what we have here is a conventional argument for a conventional proposition — to the extent that we are able to name and talk about sensations, they cannot be private. Since the argument apparently rests on the assumption that the only acceptable criterion for judging whether a word is successful in latching on to its referent is a public one, this would be a natural starting point for a counter-attack. But for a more complete understanding of the private language remarks, we need to situate them within the overall project of the *Philosophical Investigations*, which is to build up a picture of the way language functions that is quite different from that encountered in conventional philosophical discourse. Moreover, the medium of the *Philosophical Investigations* is faithful to its message. Where a conventional argument would proceed from proposition to proposition in the given language of analytic philosophy, Wittgenstein's treatment of the same topic moves from a quoted proposition expressed in the language of his interlocutor to an unquoted riposte that calls into question that very form of expression.

The overall effect is not to demonstrate that such and such conclusion follows from such and such premises, but rather to undermine a whole way of speaking, and with it a whole way of thinking. The private language remarks, along with

---

[12] Although it predates Chalmers (1996), the treatment of the zombies argument in Dennett (1991) owes much to Wittgenstein. Eilan (2002) offers a rare example of a directly Wittgenstein-inspired discussion of Chalmers, albeit in a somewhat different style to the present contribution. Eilan goes so far as to claim that 'Wittgenstein's conception, in the other minds case, of what the hard problems of realism about consciousness are, is light years ahead of where they are said to lie in much current writing on consciousness' (p. 194).

many other passages in the *Philosophical Investigations*, attack dualistic thinking by showing that it arises when language is twisted out of its everyday use to conjure up an illusory metaphysical division between inner and outer. So when Chalmers asserts that the concept of a zombie is coherent, however empirically unlikely, because he can 'discern no contradiction in the description' (Chalmers, 1996, p. 96.), he unwittingly reveals the dualistic horse lurking behind his sceptical cart.

One option for the philosopher thus confronted is to pretend that nothing has changed, and to continue to present arguments forever condemned to arrive at a dualistic destination because they proceed from a dualistic point of departure.[13] But Wittgenstein offers a radical alternative, which is to see language, not as a vehicle for communicating thoughts in which a word like 'pain' has to stand for something in the way that a word like 'brick' typically does, but rather as an activity that is embedded in the richness and complexity of our everyday human affairs (see Wittgenstein, 1958, §309). The private language remarks are an attempt to head the philosopher off at the dualistic pass, to steer him away from the ill-conceived uses of sensation language that lead to a bogus inner/outer distinction.

An exemplary exchange is contained in §309 of the *Philosophical Investigations*, in which Wittgenstein's imaginary interlocutor accuses him of repeatedly concluding that 'the sensation itself is a *nothing*'. But Wittgenstein refuses to be drawn into the dualism inherent in the phrase 'the sensation itself'. 'It is not a *something*,' he responds, 'but not a *nothing* either!' and he reiterates the point of the private language remarks: 'The conclusion was only that a nothing would serve just as well as a something about which nothing can be said.'

If Wittgenstein's strategy was merely to insist on a kind of linguistic censorship it would not be very effective. But his real aim is to promote a fundamental shift in thinking, one that results in an improvement in our epistemological predicament. This improvement is encapsulated in Wittgenstein's phrase 'nothing is hidden',[14] which is to say nothing is *metaphysically* hidden. My experiences are just as much out there as in here. When you or I dissemble, our feelings are only private in the unmysterious sense that a ball can be hidden under a magician's cup. In both cases, a more detailed enquiry would reveal all.

In taking this position we find a potential ally in McDowell, who is concerned to deflect the Cartesian thought that *this* — my inner life — might exist without *that* — the external world (McDowell, 1986; 1994).[15] The Cartesian thought mirrors the argument from the logical possibility of zombies, in which we are encouraged to accept the possibility of *that* — a living human brain — without

[13] Chalmers (1996) declines to tackle the private language remarks, making only passing reference to them in a somewhat dismissive footnote (p. 381).

[14] Wittgenstein (1958), §435. Although the phrase arises there in a rather specific context, it is often taken to be a representative slogan for Wittgenstein's later writings. Certain remarks about what is 'hidden' in Part II of the *Investigations* are especially apposite here (Wittgenstein, 1958, pp. 220–2).

[15] Indeed, McDowell's reading of the private language remarks (McDowell, 1994, pp. 18–23) is broadly compatible with the interpretation on offer here, although the Kantian vocabulary he deploys indicates a divergence of concerns.

this — an inner life. McDowell uses the case of singular thought to argue for the *openness* of our minds to the world, by which he means that our minds directly experience the world as it is, without mediation through a film of appearances. In similar terms, we might now speak of the world's openness to our minds, meaning that our experiences are directly manifest in the world.

The outcome, as far as the agenda of this article is concerned, is a licence to investigate consciousness empirically. This will address neither the easy problem nor the hard problem, for the possibility of separating these concepts no longer arises. Moreover, according to this way of thinking, an investigation into A-consciousness *is* (part of) an investigation into P-consciousness, although the extent to which it is sufficient in that role will depend on its empirical success. In lieu of a reformulation in more theoretically informed terms, we may unashamedly refer to an empirical investigation of *consciousness*.

Approval of such a licence is by no means implicit in the writings of either Wittgenstein or McDowell, in spite of their respective stands against dualism. However, consider §412 of the *Philosophical Investigations*, where Wittgenstein contrasts the *problematic* utterance of the sentence 'THIS is produced by a process in the brain!' when accompanied by 'the feeling of an unbridgeable gulf between consciousness and brain-processes', with the *unproblematic* utterance of the same sentence in the context of a hypothetical scientific experiment that might nowadays be called an investigation of the neural correlates of consciousness. Although Wittgenstein's own interests are confined to philosophical problems and their dissolution, there is nothing in his writing to suggest hostility to the idea of such an empirical programme.[16]

The same cannot perhaps be said for McDowell. Reality, for McDowell, though metaphysically undivided, is not exhausted by the 'realm of law' which is the domain of the natural sciences. It also encompasses the 'space of reasons', the logical structure of relations that is revealed by rational argument. In McDowell's view, a mature human being truly experiences the world thanks to a 'capacity to resonate to the structure of the space of reasons' (McDowell, 1994, p. 109). A 'mere animal', by contrast,

> does not weigh reasons and decide what to do [and] the milieu it lives in can be no more than a succession of problems and opportunities, constituted as such by [immediate] biological imperatives (p. 115).

So experience, under McDowell's conception, seems simultaneously to be excluded from the domain of scientific enquiry and denied to 'mere animals'.

McDowell credits non-rational animals with 'sentience' rather than experience proper, and 'proto-subjectivity' rather than fully-fledged subjectivity. However, if experience proper and fully-fledged subjectivity are construed as requiring the weighing of reasons, it is hard to avoid the unwelcome corollary that human beings ordinarily experience very little in their everyday lives (unless they are philosophers, perhaps). On the other hand, it is possible to regard the

---

[16] 'The question whether the muscles of the larynx are innervated in connection with internal speech, and similar things, may be of great interest, but not in our investigation' (Wittgenstein, 1958, p. 220).

experience of a rational human and the sentience of other animals as lying on a single continuum of consciousness. From the standpoint of the present article, the transition from mere animal to rational human looks less interesting than the transition from mere thing to conscious subject, where the latter concept is permitted to encompass both beast and philosopher.[17] Adopting McDowell's terminology, we may still grant that there is something special about the experience of a mature human being insofar as it engages more fully with the space of reasons than the sentience of a newborn baby or of another animal, while denying that this specialness significantly reduces the scope of a scientific investigation into consciousness.

4. The Subject/Object Distinction

A successful challenge to the very idea of a metaphysically divisive inner/outer distinction is enough to undercut a whole class of arguments that seek to limit the reach of a scientific enquiry into consciousness. The preceding section concentrates on the argument from the logical possibility of zombies, but the same Wittgensteinian considerations apply to all but one of the five representative arguments reviewed by Chalmers (1996, ch. 3). The exception is the 'colour scientist' argument, originally due to Jackson (1982) and often referred to as the 'knowledge argument' (Ludlow *et al.*, 2004), which makes appeal to the more generic intuitions that underpin the work of Nagel (1974; 1986).

The argument invites us to consider Mary, a leading neuroscientist some time in the future when our understanding of the brain is complete. Mary's field of expertise is colour vision, and within this domain her knowledge is total. However, Mary has lived her entire life in a black-and-white room. She has never seen colours. She does not know, for example, what it is like to see the colour red.

> No amount of reasoning from the physical facts alone will give her this knowledge. ... It follows that the facts about the subjective experience of color vision are not entailed by the physical facts. If they were, Mary could in principle come to know what it is like to see red on the basis of the physical facts. But she cannot (Chalmers, 1996, p. 103).

In a similar vein, Nagel (1974) asks us to consider the extent to which we could ever know what it is like to be a bat. Bats fly and use echolocation, so they are very different creatures to human beings. It is surely like *something* to be a bat, but no human being could ever truly know *what* it is like. Once again, the conclusion is that all scientific theories of consciousness are necessarily partial, because they cannot account for the subjective character of experience.

The thrust of this section and the next is that such arguments conflate a legitimate subject/object distinction with a bogus inner/outer distinction, causing the former to inherit the spurious air of metaphysical mystery that attaches to the latter. The subject/object distinction is intimately related to the concept of a point of view, and the contention is that this concept can be cleanly severed from the

---

[17] As Hurley says in a related context, 'we should not overintellectualize the mind' (Hurley, 2003).

already discredited inner/outer distinction and thereby rendered metaphysically harmless. This will be achieved by reconciling the concept of a point of view with an empirically supported and theoretically well-founded distinction between conscious and non-conscious information processing in the human brain, specifically that consciously processed information is globally accessible while non-conscious information processing is localised.

1. Global Access

The principal claims of Baars's global workspace theory are that (i) the human brain instantiates a particular information processing architecture, and (ii) it is possible, within that architecture, to distinguish conscious from non-conscious information flow (Baars, 1988; 1997). The architecture in question comprises a set of parallel, independent, specialist processes which interact through a global workspace (Fig. 1). The parallel, specialist processes compete for access to the global workspace, and whatever information the winning process (or coalition of processes) has to offer is thereby broadcast back to the entire set of specialists.

In addition to the separate concurrent threads of computation associated with the many parallel processes of the global workspace architecture, a distinctive and unique thread of computation is discernible in the sequence of broadcast global workspace states. This serial thread of computation is the integrated product of the parallel processing of all the specialists. The first claim of global workspace theory, then, is that the human brain instantiates a global workspace architecture of this sort. Specialist, parallel brain processes are hypothesised to be responsible for one or more aspects of perception, action selection, working memory, long term memory, language understanding and production, and so on. In addition, the theory requires the existence of some neural mechanism for arbitrating competition among coalitions of these processes and for broadcasting the message of the winning coalition.
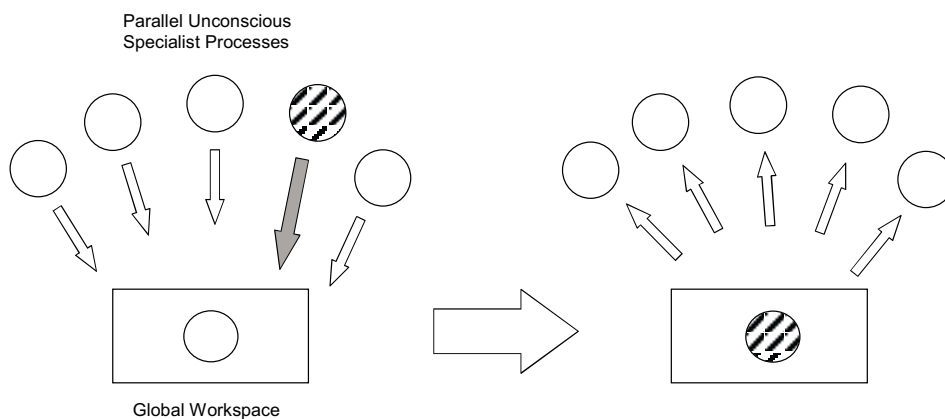

Figure 1. The Global Workspace Architecture

The second claim of global workspace theory is that information processing carried out within the parallel specialists is non-conscious, while only information that is successfully broadcast to the whole cohort is consciously processed. It is possible to investigate this claim — which only makes sense in the context of the first claim, of course — using the experimental paradigm of *contrastive analysis*. In experiments conforming to this paradigm, a conscious brain process is compared with a similar non-conscious brain process.

For example, Dehaene *et al*. (2001) used backward visual masking and fMRI to compare cerebral activity in otherwise similar cases of conscious and non-conscious processing of a visual stimulus. In the first (unmasked) case, a subject is presented with a word for a short period which is immediately followed by a blank screen. In the second case, the subject is presented with a word for the same short period but the word is immediately followed by a meaningless pattern (the mask). In the unmasked case, the subject can report the word, while in the masked case the subject typically cannot. However, in the masked case, the unreportable word still has a priming effect, thus demonstrating that it has to some extent been processed by the visual cortex.

In the results reported by Dehaene *et al*., increased brain activity occurred following the presentation of the word in both cases. But in the masked case this was localised within the visual cortex, while in the unmasked case increased brain activity was widespread throughout cortex. A considerable body of empirical evidence of a similar sort has accumulated in recent years (Baars, 2002; Baars *et al*., 2003), all of which lends provisional support to the central claim of global workspace theory that non-conscious information processing is isolated within separate specialist brain processes while information that is consciously processed is broadcast to multiple brain processes.

In addition to the empirical evidence, theoretical support for the hypothesis that the brain instantiates a global workspace architecture comes from consideration of the *frame problem* (Fodor, 1983; 2000). Fodor contends that, while some cognitive processes will fall to a computational account, others — specifically those that are *informationally unencapsulated* — are computationally infeasible. An informationally unencapsulated process is one capable of drawing on any domain of belief. Analogical reasoning is an extreme example, since by nature it is concerned with discovering relationships between domains previously thought to be unrelated.

An informationally unencapsulated cognitive process is computationally infeasible, according to Fodor, because of the unmanageably large task of sifting the relevant from the irrelevant among the myriad beliefs an agent holds. This is a manifestation of the frame problem, in Fodor's sense of that term. However, this problem only seems unmanageable in the context of a serial process that examines beliefs for relevance one at a time. In the context of a global workspace architecture, however, the challenge of bringing all and only the relevant information to bear on a problem looks less computationally forbidding (Shanahan & Baars, 2005). This is because the responsibility for determining relevance is distributed among the parallel specialist processes. Each of these processes lies

dormant until it notices that it has something of relevance to contribute to the ongoing sequence of broadcast states, at which point it wakes up and starts to lobby for access to the global workspace.

It is the contention of the present article, not only that an empirical enquiry into consciousness itself is legitimate, but also that the distinction between conscious and non-conscious information processing offered by global workspace theory could form the cornerstone of such a theory. If it were possible to demonstrate an example of consciousness without global broadcast then it would falsify the second of these contentions (but leave the first intact). One of the more subtle attempts to do this is made by Bayne and Chalmers, who describe an experiment from 1960s which they interpret as showing that it is possible to have access consciousness of state *A* and access consciousness of state *B* while lacking access consciousness of the conjoined state *A & B* (Bayne & Chalmers, 2003). Moreover, they contend that such a breakdown of unity is not possible for 'phenomenal consciousness'.

In the experiment, originally described by Sperling (1960), subjects are briefly presented with a matrix comprising three rows of four numbers each. The matrix is then removed, and the subjects are asked to recall the contents of a given row. Sperling reported that subjects were able to do this fairly reliably whichever row they were asked to report on (on average managing 3.3 correct numbers out of 4), although they were not told which row this would be until after the matrix was removed. Despite this, when subjects were asked to report the contents of the entire matrix, their performance was extremely poor, averaging only 4.5 correct numbers out of 12.

> In this case, it is natural to hold that the subject … is access-conscious of the contents of any individual row. … But it also seems that the subject is not access-conscious of the whole matrix … (Bayne & Chalmers, 2003, p. 35).

Bayne and Chalmers maintain that, by contrast, 'there is something it is like for the subject to see the entire matrix' (p. 36), and a consequence of this is that 'access consciousness and phenomenal consciousness can come apart' (p. 36). Now, we do not need to resort to the dualistic language used by Bayne and Chalmers, with its tendency to drive a metaphysical wedge between inner and outer, to see their example as a challenge to the claim that consciously processed information is necessarily broadcast. To meet this challenge, we need to pay attention to issues of time and memory that are glossed over by Bayne and Chalmers.

According to the view being championed here, consciously processed information is necessarily broadcast, and it makes no sense to speak of broadcasting information state *A* and information state *B* without also broadcasting information state *A & B*. In Bayne and Chalmers terms, this entails that conscious states are necessarily 'access-unified', a proposition they reject on the basis of cases like the Sperling experiment. But to understand the Sperling experiment properly, we must consider the information being broadcast at the time of the presentation of the matrix separately from the information being broadcast at the time

of report, and also take account of the changes that take place in the subject's working memory during the intervening period.

It is entirely possible for the layout and some details of the entire matrix to be broadcast during presentation, while the detailed image is unconsciously preserved in a short-term visual buffer, part of working memory. When asked to report a row, some of the contents of this visual buffer, corresponding to the requested row, will make its way into the global workspace for subsequent broadcast. But the visual buffer will fade rapidly, making it unlikely that the contents of other rows can reliably be given the same treatment. Moreover, the very act of reporting the contents of one row may well interfere with the unconscious specialist processes that are responsible for holding numbers in working memory.

To conclude this overview of global workspace theory, it is worth articulating a number of denials, since it is easy to misconstrue the theory in various ways. First, the global workspace itself is not a Cartesian theatre (in the sense of Dennett, 1991), a homunculus, an inner observer, or anything similarly irreducible. It is strictly an information processing concept, and its deployment in working software systems (Franklin & Graesser, 1999; Franklin, 2003) and in brain modeling experiments (Dehaene *et al*., 2003; Dehaene & Changeaux, 2005; Shanahan, 2006) demonstrate that it meets the criterion of implementability.

Second, although global workspace theory presents the consciousness of information processing as an all-or-nothing property — information is either broadcast or it is not — it should be borne in mind that (a) this is the typical, paradigmatic, non-pathological case, and (b) even in the paradigmatic case some aspects of, say, a visual stimulus can be broadcast while others are not. A vague awareness of having seen something undefined can be accounted for in terms of the wholesale broadcast of information pertaining to the presence of the stimulus plus the wholesale lack of broadcast of its shape, colour, and so on.

Third, the theory does not entail a commitment to a language of thought (in the sense of Fodor, 1975). The messages broadcast from the global workspace are, at bottom, nothing more than electrical signals with the capacity to influence the activity of the neural assemblies that receive them. When, later in the article, we refer to the 'encoding' of these messages, this merely alludes to the possibility of a discoverable systematic relationship between the underlying signals and the ongoing situation in the brain, the body, and the environment.

2. Embodiment

Despite the explanatory appeal of global workspace theory, the overarching argument of this section is that instantiating a global workspace architecture is not enough in itself to constitute a subject. Instead, the experiencing subject is seen as a *coincidence of unities*, or better, a *cascade* of unities, since they constitute a series, each member of which is entailed by its predecessor. The spatial unity of the body entails the unity of the senses' viewpoint on the world. Together, these corporeal unities entail a unity of purpose of the parallel specialists that contribute to

the global workspace, and this in turn entails the perspectival unity of the contents of the global workspace itself. In this way, the access consciousness realised through the global workspace architecture is bound to the spatially confined body it inhabits, and the result is an experiencing subject.[18]

Every biological organism can be spatially individuated according to an unbroken spatial boundary, which is to say we can imagine a closed, contiguous membrane stretched over the organism such that everything that is part of the organism is within the membrane and everything outside the membrane is foreign to the organism. This definition allows for prostheses, implants, pregnancies, parasites, and so on. It also allows for a membrane that is larger than necessary, so in what follows we will assume a criterion of spatial individuation that appeals to the smallest membrane permissible, which will be referred to as the *bounding membrane*.[19]

The bounding membranes of both plants and animals are liable to change shape considerably during their lifetimes and, in the case of an animal, with some vigour. Moreover, in the human case, we speak of the food we ingest becoming part of our body, and of any limb that we lose ceasing to be part of our body. During the course of pregnancy, the bounding membrane of a new organism forms within the bounding membrane of an older one. But none of these common-sense considerations allows for a single organism with a fragmented body — a single organism with two spatially separated bounding membranes. In this sense, we can speak of the spatial unity of the body.

Using the technology of telepresence, a person can project her point of view to a remote location. But this possibility does not conflict with the thesis under development, because such an ability is parasitic on that person's having an already established point of view that is centred on her natural body. The aim here is solely to lay down a series of physical and informatic conditions, each building on the last, that culminates in the possibility of a conscious subject. How that series might be continued further is another matter.

An organism's entire sensory apparatus is contained within its bounding membrane. This imposes a spatial constraint on the information it can deliver. In the case of a tactile sense, this constraint is very strong — the source of a stimulus has to be in contact with the organism's body. In the case of a visual sense, the source of the stimulus can be very remote from the organism's body indeed — a distant astronomical object, for example. Nevertheless, the size, position and, thanks to occlusion, the very presence of an object in the visual field all depend on the organism's location, and each of these things changes in a systematic way when the organism moves. Whatever the modality, the containment of an

---

[18] Contemporary cognitive science is, of course, replete with affirmations of the importance of embodiment. For example, Varela, *et al.* (1991), Clark (1997), and Wheeler (2005) all make the case that cognition is inherently embodied, and their arguments sit well with the present paper. However, our primary concern here is consciousness rather than cognition. Gallagher (2005) discusses the ways in which the body conditions conscious experience, but does not directly address the question under investigation here, which is whether embodiment is a necessary condition of subjectivity.

[19] Note that the bounding membrane, as defined here, is a spatial abstraction, not a physical entity.

organism's sensory apparatus within its bounding membrane ensures that every sense is dependent on the same location, the location of the organism's body.

In this way, we can speak of the unity of the senses' viewpoint on the world, a property that is inherited from the spatial unity of the body. As with the possibility of telepresence, the variety of ways humans have invented to extend the reach of their senses to remote locations, such as telephones and webcams, does not diminish the point. The possibility of such devices piggybacks on an already extant natural sensory apparatus whose locus is the spatially confined body to which it is fastened.

3. The Conscious Subject

One requirement that any instantiation of a global workspace architecture must meet is that the information processed by the set of parallel specialists must be coded in such a way as to be 'generally intelligible'. That is to say, an item of information is only worth broadcasting if it has the potential to influence usefully the activity of any of the processes that receive it, and this demands a coding scheme that can be understood by all such processes. One important property of a generally intelligible coding scheme is that every message employing it must unproblematically identify which aspect of the outside world it is about.

The general intelligibility issue is related to the so-called *binding problem* (von der Malsburg, 1999; Treisman, 1999). The distributed nature of the biological brain means that properties belonging to one and the same object are processed by different brain regions, while sets of properties that belong to quite distinct objects are processed by the same region. Suppose a person is looking at an image containing both a red square and a green triangle. The colour of the green triangle will be handled by a distinct region of visual cortex from the shape. Yet the colours of both the triangle and the square will be handled by the same region. The binding problem is to understand the neurological mechanisms that ensure that the colour red (and not green) is associated with the square (and not with the triangle).[20]

The issue at hand is related to but subtly different from the binding problem. The binding problem takes for granted an embodied biological brain and the constraints imposed by its anatomical organisation. By contrast, the question of general intelligibility is initially neutral about embodiment, but takes the global workspace architecture for granted, based on the premise that it can explicate the distinction between conscious and non-conscious information processing. If we step away from biological brains and consider the global workspace architecture in the abstract, then we can imagine a variety of forms that general intelligibility might take, depending on how much common ground the specialist processes have and how much contextual information can therefore be taken for granted when information is passed between them.

---

[20] A variety of proposals have been put forward for how the brain might solve the binding problem. Implicit in all of them is the assumption of a single, spatially confined body and a shared sensory apparatus. The binding problem would be a good deal harder if the brain processes handling colour and shape could not take for granted a single, common field of view.

For example, in the context of a genuinely fragmented body and a genuinely distributed suite of sensors — where 'genuinely' indicates a lack of parasitic dependence on pre-existing corporeal unities — it might be possible to achieve general intelligibility through the use of complex definite descriptions to pick out which aspect of an ongoing situation a broadcast message is about. To see this, let's consider the analogy of a rural community served by a local radio station. Members of the community are always on the lookout for interesting news items — village fetes, jumble sales, car accidents, and so on. When they think they have something to report, they phone their local radio station, which chooses whether or not to broadcast the information.

The combination of community and radio station in a sense instantiates a global workspace architecture, with the members of the community fulfilling the role of the parallel specialists and the radio station taking the part of the global workspace. There is sufficient common ground between these 'parallel specialists' for the general intelligibility of the messages they phone in to the radio station to be achieved through the careful use of proper names and definite descriptions. Someone who phones in to report a traffic incident cannot use phrases like 'this car' and 'over there' and 'him' to describe the situation to someone who is not on the scene. But she can refer to 'the bypass east of Aylsham' and 'the eldest son of the man who owns the King's Arms'.

In comparison to the members of this rural community, a much greater degree of common ground is enjoyed by an assembly of parallel specialist brain processes, each of which is inserted in the same overall sensorimotor loop, a sensorimotor loop that controls a single, spatially unified body using the data delivered by a sensory apparatus attached to that body. This common ground underpins an especially intimate form of general intelligibility, wherein the information flowing into and out of the parallel specialists has an indexical character.

Linguistically speaking, a term is *indexical* if its reference varies systematically with the context in which it is used. Typical indexical terms include 'here', 'now', and the demonstratives 'this' and 'that'. But the concept of indexicality can be widened to include any message whose encoding depends on the context in which it is sent. For example, the feedback path of a closed-loop control system — as used, say, to regulate the temperature of a shower — carries an error signal which the engineer might (crudely) characterise in terms of phrases such as 'too hot' or 'too cold', which implicitly make reference to 'the temperature here now'. The error signal is only intelligible given the context of the ongoing situation for the system in question. But, duly given, the context can be taken for granted and does not have to be represented in the signal.

In a similar fashion, in the given context of a spatially localised body, messages input to and output from a set of parallel specialist processes can implicitly refer to the ongoing situation from the viewpoint of that body. A message may have to explicitly encode which aspect of that ongoing situation it is about — using a spatial tag, for example. But, thanks to the common ground of a shared body and sensory apparatus, this encoding *never* has to pick out an object from the world's total inventory, and at worst it has to do so from among only a small

number of candidates. Nothing analogous to a definite description has to accompany a message to identify the object it pertains to. If a cat enters a room and draws a person's gaze, then any process in that person's brain can potentially send or receive a message that is unambiguously about *this* animal.[21]

Furthermore, for the unfolding contents of a global workspace to make sense — for it to exercise any form of coherent influence on the world — the parallel specialist processes that contribute to it must serve a common remit. Even though these processes compete for access to the global workspace, the very idea of a winner only makes sense if there is a universal criterion for success, a criterion that selects the process or coalition of processes that is, in some sense, most relevant at any given time to an ongoing situation. And in much the same way that embodiment facilitates the general intelligibility of the messages input to and output from the parallel specialists, embodiment also supplies them with a common purpose, which is to maintain the integrity and wellbeing of *this* body, the same body that fixes the viewpoint from which information about the outside world is taken in. Moreover, the only way to fulfil this duty is by moving muscles that are themselves components of the very same body.

So, when a global workspace architecture is instantiated by a biological organism, the information processing carried out by each of its parallel specialist processes is thoroughly conditioned by its indexical relationship to the organism's body — in terms of input, output, and functional role. And the very structure of the global workspace architecture entails that the indexicality of information processed non-consciously by the parallel specialists is duly inherited by consciously processed information — that which is successfully broadcast to the entire set of specialists from the global workspace. A subjective point of view — in a sense that is more than merely spatial — is thus conferred on the conscious procession of broadcast states, and thereby on the organism's consciousness itself. The result is an experiencing subject.[22]

5. Subjectivity Revisited

Let us briefly review the steps that have brought us to this conception of the experiencing subject. First, a series of critical remarks were advanced with the aim of legitimising an empirical investigation into consciousness *per se*, an

---

[21] Campbell's (2002) treatment of the interplay between consciousness, attention, and indexicality is of considerable interest here. Campbell's emphasis is on the role of conscious attention in determining 'demonstrative reference', wherein an explicit or implicit act of pointing sets the context for picking out the referred to object among many candidates. The present article gives a still more fundamental role to a different form of indexicality, one at the information processing level rather than that of language, and one that is made possible by embodiment. Campbell asks 'just how conscious experience of the object manages to identify the target of the information processing machinery' (p. 3). By contrast, the present article claims that the combination of embodiment and a global workspace architecture is what makes it possible for a piece of massively distributed information processing machinery even to have such a thing as a single target, and that this in turn makes conscious experience possible.

[22] A compatible treatment of the relationship between consciousness and embodiment is given by Legrand (2006), who emphasises the significance of the body as the 'point of convergence of perception and action'.

investigation that eschews the dualistic distinction between inner and outer. An invitation was extended to see the world as a place where nothing is hidden, metaphysically speaking, entailing that experience is just as much out there as in here. Then, in pursuit of the sort of investigation thus legitimised, endorsement was given to global workspace theory. This endorsement is provisional, of course. Specific theories are always vulnerable to empirical refutation. However, current empirical data is supportive of global workspace theory, and in addition there are *a priori* arguments in its favour.

The next step in the argument was the claim that to instantiate a global workspace architecture is not sufficient in itself to constitute a conscious subject with a point of view. Yet if a biological organism instantiates a global workspace architecture, a point of view is conferred on the procession of conscious content passing through its global workspace thanks to the indexical character of the information processed by its parallel specialists. This indexical character arises because all the processes share a single, spatially localised body and operate on data from a sensory apparatus that is attached to that body. The overall conclusion is that a conscious subject is a cascade of unities, starting with the spatial unity of the body and that of the viewpoint of the body's senses, leading to the unity of perspective of the parallel specialists, giving rise to the indexical unity of the procession of contents of the global workspace.

The concept of a conscious subject we have arrived at is not to be confused with that of a self, in the grand sense of that concept which throws up complex issues of personal identity, autobiography, memory, personality, and so on (Strawson, 1997).[23] Mercifully, all of these issues can be sidestepped here. According to the present view, a conscious subject can be conceived of without a self in the grand sense.[24] No doubt many species of animal lack autobiographical memory and arguably therefore fall into this category. It might be also argued that the biographies of newborn human babies are too slight for them properly to have acquired such selfhood, and that patients with severe cases of hypocampal amnesia have tragically lost theirs, although under the present conception both can constitute conscious subjects.[25]

---

[23] The word 'self' is used in a variety of senses in philosophical discussion. Obviously if someone chose to use some such phrase as, say, 'minimal self' for what is here termed a 'subject', then our only difference would be terminological. The earliest account of the self in the context of global workspace theory is Baars (1988, ch. 9). Metzinger's (2003) treatment of the self is also pertinent, since he takes 'global availability', meaning something akin to the present concept of global access, to be a vital constraint on conscious experience. One of the main thrusts of Metzinger's treatment is a challenge to the ontological status of the concept of the self, which, according to Metzinger, is not a thing or a property but the illusory artefact of a process involving the transparent deployment of a self-model (transparent in the sense that the deployed model does not see itself as a mere model). This is not incompatible with the view being developed here. Unlike the present paper, however, Metzinger's analysis of subjectivity places no particular emphasis on embodiment and its consequences.

[24] The distinction between (mere) conscious subject and self being drawn here bears comparison with Damasio's distinction between core self and autobiographical self and his related distinction between core consciousness and extended consciousness (Damasio, 2000, Chapters 6 & 7).

[25] A more imaginative dissociation of these two concepts is suggested by a thought experiment due to Campbell (1994), who asks us to consider 'a creature that, though intelligent, is like the amoeba in

Conversely, it could be argued that when someone is asleep, there is no conscious subject present, although the self persists, along with the potential for it to re-engage in conscious subjectivity. It follows from all this, of course, that self-awareness, that is to say consciousness of the self (whatever that may be), is at a still further remove from our present concerns. However, we might suppose, along with Proust,[26] that in the first few moments of re-engagement with conscious subjectivity after sleep both self, in the grand sense, and conscious subject are fleetingly present before self-awareness has had time to fully kick in.

Now, armed with this conception of a conscious subject, let us revisit Mary the colour scientist, and instead of asking what she 'learns' or comes to 'know' when she sees red for the first time, let us ask more simply what it is about her that *changes*. The bottom line is that prior to seeing red for the first time, Mary has never occupied the point of view of someone seeing red, while afterwards she has. But her occupation of such a point of view means much more, under the present conception of a conscious subject, than simply her looking in the direction of a red object. It means also that the redness of the red object gives rise to a certain pattern of neuronal activity in her visual cortex and, crucially, that this activation pattern systematically influences — is broadcast to — a whole gamut of other brain processes, which are responsible for action selection, affect, working memory, autobiographical memory, and so on.

Moreover, the visual sensory apparatus from which the broadcast pattern originates is attached to Mary's spatially localised body, and the common remit of all the brain processes receiving the broadcast is to maintain the comfort and well-being of that body given the situation apparent from the viewpoint of that sensory apparatus, which is why they can all usefully absorb a novel piece of information (corresponding to redness) that pertains to *that* object. The only distinctive thing about Mary is that this particular combination of facts has never previously been applicable to her, and this has no relation whatsoever to her own scientific understanding of facts of this sort.

Let us now turn our attention back to Nagel and his bat. Do we have a concept of the conscious subject that will enable us, at least in principle, to determine what, if anything, it is like to be a bat, or indeed an alien, or a robot? The answer is yes. We can answer the question of whether or not it is like anything at all to be an *X* (where the *X* might be a bat, an alien, or a robot) by examining the

---

that it frequently fissions and like some types of particle in that it frequently undergoes fusion' (p. 96). Moreover, we are asked to suppose that both the products of a fission 'inherit all the psychological properties of the original,' while in fusion, 'as much as possible of the psychological lives of the originals are passed on to the successor' (p. 96). Campbell concludes that such creatures cannot be persons on the grounds that their biographies are too fractured to support first-personal thoughts. For persons here, we might read 'selves' and concur with Campbell. But the current conception, allows for the possibility that such creatures would, given suitably organised nervous systems, still constitute conscious subjects.

[26] '… when I awoke in the middle of the night, … I had only the most rudimentary sense of existence, such as might lurk and flicker in the depths of an animal's consciousness'. For Proust, self-awareness is bound up with memory, which 'would come like a rope let down from heaven to draw me up out of the abyss of non-being' and allow him to 'gradually piece together the original components of my ego' (*A la Recherche du Temps Perdu*, vol. 1).

mechanisms that mediate between an *X*'s sensory input and its motor output, with a view to discovering some instantiation of a global workspace architecture. Having established that it is indeed like *something* to be an *X*, we can approach the question of *what* it is like by studying the kind of information that passes through the global workspace thus identified, a task that requires a degree of understanding of the bodies of *X*s and their sensory apparatus.

But what of the objection that this will never tell us what it is *really* like to be a bat — that we will never know what it is like for a bat to be a bat, as Nagel put it? This sort of objection can now be harmlessly reformulated. It amounts to nothing more than the metaphysically unmysterious assertion that, because the body and senses of a bat are so different from those of a human, the information that passes through the global workspace in a human brain will always have a very different character from that passing through the global workspace in a bat's brain (assuming such a thing exists). In other words, there is a certain collection of facts that will never be applicable to a human being, but this in no way limits the scope of a scientific enquiry into facts of that sort.

Acknowledgments

References

Aleksander, I. (2005), *The World in my Mind, My Mind in the World* (Exeter: Imprint Academic).

Baars, B.J. (1988), *A Cognitive Theory of Consciousness* (Cambridge: Cambridge University Press).

Baars, B.J. (1997), *In the Theater of Consciousness: The Workspace of the Mind* (New York: Oxford University Press).

Baars, B.J. (2002), 'The conscious access hypothesis: origins and recent evidence', *Trends in Cognitive Science*, **6** (1), pp. 47–52.

Baars, B.J., Ramsøy, T.Z. & Laureys, S. (2003), 'Brain, consciousness and the observing self', *Trends in Neurosciences*, **26** (12), pp. 671–5.

Bayne, T. & Chalmers, D. (2003), 'What is unity of consciousness?', in A.Cleermans (ed.), *The Unity of Consciousness: Binding, Integration, and Dissociation* (Oxford: Oxford University Press).

Block, N. (1995), 'On a confusion about a function of consciousness', *Behavioral and Brain Sciences*, **18**, pp. 227–87.

Block, N. (2005), 'Two neural correlates of consciousness', *Trends in Cognitive Science*, **9** (2), pp. 46–52.

Block, N., Flanagan, O. & Güzeldere, G. (1997), *The Nature of Consciousness* (Cambridge, MA: MIT Press).

Campbell, J. (1994), *Past, Space, and Self* (Cambridge, MA: MIT Press).

Campbell, J. (2002), *Reference and Consciousness* (Oxford: Oxford University Press).

Chalmers, D. (1996), *The Conscious Mind: In Search of a Fundamental Theory* (New York: Oxford University Press).

Clark, A. (1997), *Being There: Putting Brain, Body, and World Together Again* (Cambridge, MA: MIT Press).

Damasio, A.R. (2000), *The Feeling of What Happens: Body, Emotion and the Making of Consciousness* (London: Vintage).

Dehaene, S. & Changeaux, J.-P. (2005), 'Ongoing spontaneous activity controls access to consciousness: A neuronal model for inattentional blindness', *Public Library of Science Biology*, **3** (5), p. e141.

Dehaene, S., Naccache, L., Cohen, L., Bihan, D.L., Mangin, J.F., Poline, J.B. & Riviere, D. (2001), 'Cerebral mechanisms of word masking and unconscious repetition priming, *Nature Neuroscience*, **4**, pp. 752–8.

Dehaene, S., Sergent, C. & Changeux, J.-P. (2003), 'A neuronal network model linking subjective reports and objective physiological data during conscious perception', *Proceedings of the National Academy of Science*, **100** (14), pp. 8520–5.

Dennett, D. (1991), *Consciousness Explained* (London: Penguin edition 1993).

Dennett, D. (1997), 'The path not taken', in N.Block, O.Flanagan & G.Güzeldere (eds.), *The Nature of Consciousness* (Cambridge, MA: MIT Press).

Eilan, N. (1995), 'Consciousness and the self', in J.L.Bermúdez, A.Marcel & N.Eilan (eds.), *The Body and the Self* (Cambridge, MA: MIT Press).

Eilan, N. (2002), 'The reality of consciousness', in D.Charles & W.Child (eds.), *Wittgensteinian Themes: Essays in Honour of David Pears* (Oxford: Oxford University Press).

Evans, G. (1982), *The Varieties of Reference* (Oxford: Oxford University Press).

Fodor, J. (1975), *The Language of Thought* (Brighton: Harvester).

Fodor, J.A. (1983), *The Modularity of Mind* (Cambridge, MA: MIT Press).

Fodor, J.A. (2000), *The Mind Doesn't Work That Way* (Cambridge, MA: MIT Press).

Franklin, S. (2003), 'IDA: A conscious artifact?', *Journal of Consciousness Studies*, **10** (4–5), pp. 47–66.

Franklin, S. & Graesser, A. (1999), 'A software agent model of consciousness', *Consciousness and Cognition*, **8**, pp. 285–301.

Gallagher, S. (2005), *How the Body Shapes the Mind* (New York: Oxford University Press).

Hurley, S. (2003), 'Animal action in the space of reasons', *Mind and Language*, **18** (3), pp. 231–56.

Jackson, F. (1982), 'Epiphenomenal qualia', *Philosophical Quarterly*, **32**, pp. 127–36.

Legrand, D. (2006), 'The bodily self: The sensori-motor roots of pre-reflexive self-consciousness', *Phenomenology and the Cognitive Sciences*, to appear.

Ludlow, P., Nagasawa, Y. & Stoljar, D. (ed. 2004), *There's Something About Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument* (Cambridge, MA: MIT Press).

Marcel, A.J. (2003), 'Introspective report: trust, self-knowledge and science', *Journal of Consciousness Studies*, **10** (9–10), pp. 167–86.

McDowell, J. (1986), 'Singular thought and the extent of inner space', in P.Pettit & J.McDowell (eds.), *Subject, Thought, and Context* (Oxford: Oxford University Press).

McDowell, J. (1994), *Mind and World* (Cambridge, MA: Harvard University Press).

Metzinger, T. (2003), *Being No One: The Self-Model Theory of Subjectivity* (Cambridge, MA: MIT Press).

Nagel, T. (1974), 'What is it like to be a bat?', *Philosophical Review*, **83** (4), pp. 435–50.

Nagel, T. (1986), *The View from Nowhere* (Oxford: Oxford University Press).

Shallice, T. (1988), 'Information-processing models of consciousness: Possibilities and problems', in A.J.Marcel & E.Bisiach (eds.), *Consciousness in Contemporary Science* (Oxford: OUP).

Shanahan, M.P. (2006), 'A cognitive architecture that combines internal simulation with a global workspace', *Consciousness & Cognition*, to appear.

Shanahan, M.P. & Baars, B.J. (2005), 'Applying global workspace theory to the frame problem', *Cognition*, in press.

Shear, J. (ed. 1997), *Explaining Consciousness: The Hard Problem* (Cambridge, MA: MIT Press).

Sperling, G. (1960). The Information Available in Brief Visual Presentations. *Psychological Monographs* 498, 1–29.

Strawson, G. (1997), '"The Self"', *Journal of Consciousness Studies*, **4** (5–6), pp. 405–28.

Treisman, A. (1999), 'Solutions to the binding problem: progress through controversy and convergence', *Neuron*, **24**, pp. 105–10.

Varela, F., Thompson, E. & Rosch, E. (1991), *The Embodied Mind: Cognitive Science and Human Experience* (Cambridge, MA: MIT Press).

Von der Malsburg, C. (1999), 'The what and why of binding: A modeler's perspective, *Neuron*, **25**, pp. 95–104.

Wheeler, M. (2005), *Reconstructing the Cognitive World* (Cambridge, MA: MIT Press).

Wittgenstein, L. (1958), *Philosophical Investigations*, tr. G.E.M.Anscombe (Oxford: Blackwell).