A Final Solution to the Mind-Body Problem

John G. Taylor

Department of Mathematics, King's College
Strand, London WC2R2LS, UK *Email: john.g.taylor@kcl.ac.uk*

Abstract
Mind and body appear on first view to be very distinct entities, but yet they interact most strongly in some as yet unknown manner. The nature of these two entities and their interaction poses the mind-body problem. Here we consider a simple solution to the mind-body problem – that of physical reductionism - 'all is physical'. We present an argument that questions the very existence of immaterial entities, and that all could be composed of such. With purely physical entities to play with, on the other hand, we present a solution as to how the mind could be created from these purely physical entities by means of attention focused in the brain on a relevant target in the external world. Further extension of attention understanding leads to the CODAM model of consciousness, supported by both EEG/MEG experiments and inner experience as described by Husserl and many others. We conclude with a discussion of the nature of the ownership of our experience and of further experiments needed to justify our proposed solution. The Mind-Body problem is discussed from the idealist, dualist and completely physical aspects. Only the last of these appears to avoid a number of criticisms, but has the deep difficulty of achieving a model of the mind from purely physical activity. We propose as a solution to this problem an attention-based corollary discharge of attention approach, which both has experimental support and provides an explanation of inner ownership not derivable from other theories. We conclude that this approach provides a final solution to the age-old mind-body problem.

1. The Mind-Body Problem

The mind is composed of mental fragments - sensations, feelings, thoughts, imaginations, all flowing now in an ordered sequence, now in a chaotic fashion. There are also non-conscious components involved in early brain processing of stimuli (as in lower level processing in vision, such as in V1) or in emotions not yet in consciousness. On the other hand the body is constructed under the underlying laws of physics, and its components obey the well-enumerated laws of physiology. It is these characteristic differences between these two - between mind and body - that lead to the Mind-Body problem.

The Mind-Body problem has been in existence for several thousand years - going back to Plato, Aristotle, The Buddha and many other ancient Greek and Eastern thinkers. The problem is simple to state: the mind and the body seem to be entities of very different kinds (as just described), so how do they interact so as to produce in a person a mind able to have effects on their body (as when the person wills the body to perform some act), whilst also their body can affect their mind (as in the experience of pain)? Although the problem is simple it has as yet no satisfactory solution, in spite of the amount of time and thought devoted to it over the ages.

Much has been written on the variety of solutions to the Mind-Body problem. There are the dualist solutions (mind and body are distinct, although then the problem of how they interact becomes even more embarrassing, and is so far unresolved in any satisfactory manner); the idealist solution (there is only mind, with body/matter being merely a manifestation of mind). How mind could have created the beautiful subtlety of the unification of the nuclear, electromagnetic and radioactive forces of nature, with a proposed extension, yet unproven, to fusion of those forces with the force of gravity through superstrings, seems impossible to comprehend. A third position is that there is only body, so this is a purely reductive physical approach (but then the problem of how our amazing mental world, full of our 'mental' experiences are thereby created from matter, has to be faced). There are also nuanced versions of one or other of these three initial positions for solving the Mind-Body problem, some having great sophistication and subtlety. But none has yet gained universal acceptance.

There is also the question of what exactly is the mind? It is certainly composed of conscious components, but it would also seem to contain non-conscious ones as noted earlier, such as unconscious emotions and low-level processing as yet out of consciousness; automatic motor responses are also below the radar of consciousness. These unconscious or pre-conscious components are not problematic since they do not possess apparently non-material private components like those that conscious components appear to do. A stone is not considered as having such private experiences as do we; it's responses to action on it can be described in purely physical terms. Similarly the non-conscious processes in our minds can be accepted as arising solely from suitable brain processing, so more easily understood as components of the body (especially with the great advances in brain science tracking down and modeling in detail this pre-conscious dynamic neural activity). These non-conscious components of mind do not have (by definition) any conscious component, so there is no difficulty in expecting them to arise solely from brain activity: they can therefore be seen as on the body side of the mind-body duo. Thus the truly inexplicable part of the mind - that part to which the epithet Mind-Body problem most closely applies - appears to be that of consciousness. That position is appropriate to take here because it addresses what appears to be the most difficult part of the overall problem of mind-body interaction: how conscious experience can interpenetrate and fuse with bodily activity so that they affect each other in the way mentioned in the first paragraph. Detailed neural models of the other (non-conscious) components of brain activity are increasingly convincing such as for predictive value coded by dopamine, for motor responses guided by internal motor models, and for early models of visual and other sensory processing through a hierarchy of increasingly complex feature detectors; these brain-based activities clearly support the lack of any non-conscious mind-body gap. These features support the approach used here, and in line with modern thinking on the Mind-Body problem, especially that brought to the fore by Descartes, and since emphasized in the notions of the hard problem (Chalmers, 1996) and the explanatory gap (Levine, 1983): it is the gap between consciousness and brain activity that now constitutes the Mind-Body Problem. This paper proposes to give a simple explanation of that gap and how it can be closed.

We start in section 2 with the 'new knowledge base', in which the nature of the mind-body gap needs to be addressed anew by taking account of the many new advances in the physical and brain sciences. In the following section 3 we question the possibility of any analysis of immaterial entities, necessary for the existence of an immaterial component to existence. Further analysis of the structure of mind is detailed in section 4, with a similar updating of our analysis of matter in section 5. In the latter section we present some postulated global principles of action of the brain, which are extended to basic principles for the mind in the following section 6. Recent brain-based solutions to the mind-body problem are briefly surveyed in section 7. Having singled out the CODAM attention control model as the most promising of these approaches, we present in section 8 the development of a control approach to attention, based especially on recent fMRI and single cell data, leading to an explicit form of the CODAM model. In section 9 a formal theory of the inner self is developed, based on the results and CODAM model of section8. Experimental support for the existence of a corollary discharge component of attention from EEG and MEG data is surveyed in section 10, and the paper concludes with a brief description, in section 11, of the nature of ownership as seen from CODAM, as our simple solution to the Mind-Body problem.

2. The New Knowledge Base

The knowledge base we can bring to attempt to solve the mind-body problem has changed enormously over the thousands of years of the problem's existence. Matter has been probed down to very short distances through the development of ever higher energy particle accelerators, so that we now have an understanding of matter in terms of the gauge theories unifying the electromagnetic and weak forces and further fusion with the strong nuclear forces by means of the quark-gluon gauge theory. Together with gravity, understood through Einstein's general relativity, the comprehension of matter has improved down to distances of about seventeen orders of magnitude. Recently this distance has decreased even more by the coming on-stream of the large hadron collider, the LHC, at CERN. Hopefully it will find support for the basic features of the present level of unification, in terms of the discovery of the Higgs boson (as the basis for mass in the Universe) as well as providing us with new clues as to the next steps of probing ever deeper into the material world, such as by finding supersymmetric partners (or even supergravitational partners) to the present spectrum of *elementary* particles (for which there are even now claimed *bumps* in particle processes best explained, it has been claimed by their discoverers, by these exotic possibilities).

This progressive and impressive unlocking of the secrets of matter has only recently begun to be caught up by new ideas and new tools in brain science. These have allowed an ever better understanding of the manner in which the brain creates and controls experience. This knowledge is allowing ever deeper appreciation of how defects in brain processing brought about by various sorts of damage to the brain can enormously modify the overall experience of the person. There is still a large distance to travel to bring the knowledge of the brain up to the level of our enormously more detailed understanding of matter. But that gap appears to be closing fast, with a possible hiatus in the physics of the elementary particles due to the presently intractable character of the mathematics of superstrings (the only way known to sensibly unify gravity with the other forces of matter, Taylor et al, 1992); no such hiatus appears in brain science, except, of course, for the mind-body problem. Will that be the rock on which neuroscience founders in a similar way, it is claimed by some physicists, as that of unifying the forces of nature will similarly founder on the rock of full unification of the forces of nature by inclusion of gravity with the other three forces?

The tools of brain science have been put to good use in probing the brain both locally and globally. The machines of PET and fMRI, measuring changes in blood flow in local regions of the brain due to neural activity, and EEG and MEG, measuring the electric and magnetic activities related more directly to neural activity in the brain, are allowing an increasingly precise view of the networks active across a range of functions. Combined with trans-cranial magnetic stimulation, the timing of the flow of crucial activity across the functional networks in the brain is beginning to give up the brain’s secrets. At the same time correlated activities of small groups of single neurons in different regions of the brain are indicating how these various regions process activity at a local level, but also as part of a larger network of such regions in interaction. Ideas from dynamical systems theory have begun to help us understand better the brain at this intermediate mesoscopic level. In all, brains are becoming ever more understandable across all scales.

The major problem arising from this for the various idealist solutions outlined in the previous section to the mind-body problem are now clear: how can such precise physicalist explanations as have now been developed of the physical world at both micro and global scale be obtained from a purely mental universe? Mind does not enter into any of the detailed physical models (such as Einstein's Theory of General Relativity, or the Standard Model of elementary particles based on quantum field theory, with the increasingly well-explored and experimentally justified SU(3) x SU(2) x U(1) symmetry). The problem is as to how to derive such well-established theories from a mental approach to the universe. As far as this author knows, there is not even a single framework available to attempt to solve this problem, nor a single idealist paper discussing it.

3. Defining a Non-Material Entity

There is a further very important problem about the mind side of the Mind-Body problem which the enormous advances of high-energy physics and brain science, described above, have exacerbated, weighing the scales ever more heavily on the side of matter: How can any scientific discussion ever be given of an entity which is non-material? This question is not one to be taken lightly. If it is responded to that science is in any case not the relevant arbiter here, then the whole discussion is going to be unable to make any forward progress to resolve any of the important issues:

How can mind and matter interact in any dualistic theory. The difficulty here is that there are no comparable features of mind to those of matter. Thus where are quantitative descriptors such as energy, particles, waves, or related quantities present in detail in descriptions of matter, down to quantum fields describing the creation and annihilation of multiplets of the most elementary of particles. How to match these up to corresponding descriptors of immaterial entities conjectured as composing mind?

How could such a theory be tested in any quantitative manner if there is no such quantitative theory?

Only by appeal to the world of experience can we conclude an argument between opposing positions. If it is claimed that part of that experience is in any case purely mental, so above the scientific arena, we can only point to the powerful developments in brain science that depend on inner report of subjects but ties them down to the activities of specific parts of the brain (or their absence in the case of deficit studies). Thus inner experience is used in a scientific manner here. Even the distortions of schizophrenics and those with other mental diseases are presently being used to probe those brain components, such as the attention system, which are damaged in the diseases, and may help explain the distorted experience reported by sufferers.

At this point it is relevant to attempt to define a non-material entity. The entity must be such that it has no material features, or at least not all its features are required to be material. But that leads us to need to define a non-material feature.

We can attack the problem of defining a non-material feature in a negative manner. Thus a non-material feature is one that cannot be considered as existing in the material universe: it can have no space or time dimension or possess any energy or weight. More completely it cannot be acted on by any of the forces of nature. This definition has clarity, but is it possible to turn it into a positive definition in terms of the presence of some basic ‘immaterial attributes’? On further reflection this appears very difficult, for any non-material attribute has no possible description in present language. Thus it would appear we cannot even begin to define a non-material entity.

At this point it is usual for the interlocutor to step back and say that science has not observed all possible aspects of the Universe. 'There are more things under heaven and earth,... etc, etc'. But there are not – that claim is completely untrue - if these entities are of a material nature. The properties of matter – its decomposition into ever smaller components down to quarks, gluons and W & Z particles - all fits together so remarkably accurately down to those incredibly short distances of seventeen orders of magnitude mentioned earlier that there can be nothing material that has been missed. Granted dark matter and dark energy are still not understood, although they could fit well into the new zoo of particles I mentioned earlier that may be seen at the LHC, and in any case were involved at a very early stage of the Universe. They are not relevant to us here and now at the sorts of energies and sizes of objects we can normally deal with.

The simplest aspect of the problem of even discussing a non-material entity or feature is that it could possess no possible attributes enabling it to be compared to anything in the material world. It has no shape or length, no weight, no color, or any other attribute that can be compared with any physical entity. In other words it can only be made of indescribable attributes. It could be claimed that it must therefore be somehow composed of purely abstract attributes. But what could these be? Even purely mathematical symbols must have some quantitative ground to be made sense of. Moreover one would be falling into a Platonic trap by assuming that a mathematically defined entity actually exists in the world. It only exists in the brains of mathematicians who have worked on it. But that existence is in terms of the geometrical or algebraic objects codes used by those mathematicians to think about their concepts.

Abstract concepts such as love (used much in attributions to God in religions across the world) is itself suspect as to its meaning. If it is to be used in relation to our human interpretation then it is quantifiable (I love you very much), due to its basis in bodily processes and behavioral responses to the loved one. It is thus not immaterially based. Recent developments in appraisal approaches to emotions and how these appraisals might be created by suitable brain activity (Taylor & Korsten, 2010, along with many other researchers, some referenced there), indicate strong support for a brain-basis for all human emotions.

If an immaterial entity has no relation to any material entity then the names 'entity' or 'immaterial 'do not help. It can be claimed by the proponent of the existence of immaterial entities that such entities exist but we can never have any knowledge of them- as in the claim that 'God is ineffable'. But then we end up with attributes which can only be defined outside our experienced universe, so completely outside our ken. If that is so, why should they even be said to exist. For the claim of ineffability, of existence always outside our ken is one which gives us no information at all. The immaterial entity is forever unknown.

Nor can so-called supernatural entities fill the bill. My own studies in the arena of the supernatural (Taylor, 1980) indicates there no such entities which can be detected by modern experimental techniques. All the experiments reported in (Taylor, 1980) indicated that there were either no effects (except those brought about by conscious or unconscious cheating) or that there were those that could be explained by relatively simple normal physical laws (electrostatics, neuro-physiology, etc). Thus the supernatural avenue is a dead end.

Thus it is difficult to analyze any non-material entity in any manner. Given that the entity cannot be described in any quantitative manner at all, nothing can be said about it that can be agreed or disagreed about with any possibility of truth (where by the latter I mean by scientifically checking up on the statement). Non-material entities or components are thus very slippery ones which twist and turn between your fingers so strongly that they escape you before you have had a time to investigate them in any reportable manner. But if you cannot report on the nature of these entities then you cannot even usefully discuss them. Even with yourself. But if you cannot do even that then such entities cannot be part of any useful conversation (with others or with yourself): no progress can ever be made in probing them.

It might be said that there must be something outside this material universe, since it itself began only some 13 or so billion years ago in a big bang. What was present before that? It can't have been anything material, since the material universe did not exist before then. Could the universe have been created somehow from an immaterial entity?

This immateriality could have existed in the multi-dimensional universe that may have been present before 13 billion years ago (as in a superstring universe in 10 or 11 dimensions). But then it would not actually have been properly immaterial. We know too little of the physics of the whole universe, however, to be able to extrapolate effectively before the first 10^{-40} of a second. So we cannot presently discuss that aspect in any scientific detail. But in any case those very early times (pre-10^{-40} of a second) would still be in the material universe, so the claim of ignorance is not going to help us bring in immaterial entities in interaction with the material universe. No model exists to help us understand how there could be a phase transition, say, from immateriality to materiality just at the time of the Big Bang. So we cannot bring in immateriality by the back door, so to speak, into the material universe.

The overall conclusion of this is that idealism and dualism face difficult problems of definition of the basic components – immaterial entities - which presently have not been solved. We thus turn to physicalism, which has its own hard problems, and most particularly the hard problem of providing a detailed explanation of how mind can be created by suitably subtle dynamics of matter (Chalmers, 1996) and of crossing that seemingly unbridgeable gap between mind and matter (Levine, 1983). This is still unsolved in a universally accepted manner, but various promising approaches have been developed over the last few decades. These involve the reduction of the mind to the dynamics of brain activity involving some, if not all, of the billions of nerve cells in the brain. We will consider this at a functional level, so do not explore building up to that from the lower single nerve cell level (which however is increasingly effectively being achieved). So we turn to look at possible global principles for the brain - its crucial functional components.

4. Further Analysis of the Mind

In the previous section we arrived at the disturbing conclusion that 'immaterial mind' is not a concept which can be grounded in any known scientific manner, so is outside the universe of scientific discourse. Since several billion people on the earth would seem to have solved the problem of talking about mind, through a variety of religious faiths, we cannot treat the result of the previous section without some further discussion.

In the previous section it was shown that 'mind', regarded as the immaterial component of our own experience, would never seem to be a concept available for scientific investigation. That is a conclusion to be expected, since we define science as a method to analyze the material world around us, and definitely not the non-material world, what ever that is. But what is this supposed further component of the universe? Let us turn to consider how it could be approached.

Our main thesis in this paper is that we can solve the Mind-Body Problem in a simple manner, through a material reductionist method: all is matter. But such a claim would seem to be immediately countered by that just stated: the mind will never be available for scientific investigation. To resolve this dilemma we need to expand the notion of science to include some aspects of personal experience. In any case that has already been done in many neuropsychological experiments, where inner report is crucial to understand the results. In fact in many such experiments such inner report is part of the 'scientific' results. So the notion of science has thereby been expanded by default.

As an example of such an expansion, consider an experiment claimed to be testing the relation between attention and consciousness (Rahnev et al, 2009). In it a subject tries to detect two sets of stimuli: those at the centre of the field of view and those in a peripheral position. In order for the central viewing to be paid most attention, the task there is a very difficult one, such as detecting the presence or absence of a dimly illuminated shape. Then the simultaneous detection of a peripheral figure will be a difficult additional burden on attention. Yet results seem to indicate that such a burden can be handled by consciousness of the peripheral figure. So apparently consciousness and attention have thereby been separated by such an experimental paradigm. But crucially the experiences of the subject are being taken into account as part of the results of the experiment.

Many more experiments have used this methodology to explore the manner in which brain activity is related to mental experience reported by the subject. Thus the attentional blink is another important experimental paradigm, in which a subject views a rapid serial presentation of visual stimuli (suc as letters or numbers. Their task is to detect two targets: target 1 (T1) is to be detected, to be reported on later, and then subsequently a second target (T2) is to be detected. The time interval between T1 and T2 is varied. From subject report it is found that there is a maximal loss of detectability for T2 if the time difference between T1 and T2 is about 300 msecs. When this occurs it has been suggested that attention has 'blinked its eye shut', hence the name 'attentional blink' (REFS). In the case of a recent study of the brain basis of the attentional blink there is a clear dependence on the subject's experience, they being asked to rate their level of awareness of the second target T2 (Sergent et al, 2005); this level and related brain activity in certain cortical areas fit well with a recent model of the attentional blink phenomenon (Fragopanagos et al, 2005).

The above paradigms, and many more, show that such an expansion of the notions of science as purely objective to include inner private experience has occurred, then further account of inner experience must be allowed. Where this stops is unclear, although experiments taking account of numerous subjects, not just depending on the claims of a single individual, would seem to have the best chance of being veridical; the accounts of single individuals, such as those with blind-sight, are good examples as to where controversies can arise due to the actual experiences of subjects and how they are best to be interpreted in an object view.

Even if we accept much in the nature of inner report, we still have to go back to the brain and its dynamical activity in order to understand, from the physicalist position, how such inner report arises in the first place. So we need to consider if we can discover general principles of how the brain might function to produce such inner report.

5. Global Principles of the Brain's Action

An important part of understanding how reportable activity could arise purely from brain activity is that of the extraction of basic principles of brain dynamics at a global level to help comprehend the awesome power possessed by these coupled circuits of local nerve cells. Several global principles can be extracted from the welter of possibilities. Three of these are arguably the most important:

Attention, as a control system to filter lower-level brain activity so as to allow very few stimulus representations to enter the higher-level arena of thought and manipulation of neural activities; these filters are mainly controlled by activity in parietal and pre-frontal cortices. Attention thereby allows the higher level processes such as thinking and reasoning to work on a much smaller number of stimulus representations. Without such filters we would be swamped by the 'blooming buzzing confusion of the external world', as William James put it;

Emotion, in terms of value maps learnt in orbito-frontal cortex (and also coded in associated amygdala sites) so as to bias what is to be processed and to guide choice of task goals (by their associated predicted rewards). These value maps, fused with body activations and automatic brain-stem responses, are used to give emotional value to decisions for action;

Long-term memory created on-line, so as to allow for incremental wisdom about the environment for use as a guide for further interactions.

Other global principles can be added to these three possibilities, such as the use of hierarchical processing (as noted especially in vision, so as to create flexible visual codes for complex objects which can be used at a variety of scales). Another possibility involves the use of synchronization of neurons over long distances, such as at 40Hz and other frequencies, so as to solve the binding problem of combining the different codes for objects (as occur in multi-modal hierarchical coding schemes described earlier). A further principle is that of recurrent loops of neural activity, to allow for the creation of short term memory (working memory) sites for the temporary holding of such activity for spreading around to other similar sites, so acting as 'report' centers in the brain. Finally there are the principles of integration and segregation of the neural systems of the brain, as discussed for example, in (Moutoussis, 2002), which also play an observable crucial role in brain processing efficiency.

6. Global principles of the mind's action

In order to make progress on how the conscious mind is created by brain activity, we must also outline what it is we consider is the main structure of the mind. This is a highly controversial area, with many possible suggestions. However a certain degree of clarity seems to have arisen from recent ideas about the bipartite nature of consciousness by Western phenomenologists (Zahavi, 1999, 2005b, 2006). From this increasingly influential body of writing, consciousness can be seen to be divisible into 1) Content. This is available for report and internal manipulation, being composed of the neural activities representing the colors, hardness and so on of objects, for example of the specific scent of the rose and the taste of the wine, and more generally of the various modality-based features of external objects in our consciousness; 2) Ownership. This is termed by Western phenomenologists the pre-reflective or inner self. The inner self is to be regarded as that denoted by the singular I. The inner self is supposed to have no content, the term pre-reflective implying that it is at a more primitive level than any reflective descriptor of oneself, such as possessing a beard or moustache or being blond. The pre-reflective self is required to be present in the mind as an independent entity along with content in order to prevent an infinite regress which would otherwise arise when considering I as defined by reflection on one's characteristics (Shoemaker, 1968). Thus although the pre-reflective self is content free it is an essential companion of the content of experience: without an owner there can be no content, no-one to have the experience. This aspect has been especially emphasised by the French philosopher Merleau-Ponty (Merleau-Ponty, 1945) where he noted that the word consciousness has no meaning independently of the pre-reflective self. Sartre (Sartre, 1943) also pointed out that the pre-reflective self constitutes the very mode of being of experience. It provides the bedrock of 'what it is like to be' in terms of ownership of the contents of experience; I am the experiencer (Nagel, 1974). Such a view is also supported by Western philosophers such as (Flanagan, 1992) who points out that a form of low-level consciousness (identifiable with the pre-reflective self) involves crucially experiencing an experience as mine (so as being the owner of the experience). A further aspect of the pre-reflective self accepted by these thinkers is that it is a constant and immediate component of consciousness, present "whenever I am living through or undergoing an experience" (Zahavi, 2005b, p1; ibid, 2006).

We should note here that this extra component of consciousness, the pre-reflective self, was missed by the influential British philosopher David Hume (Hume, 1739) who searched for it in his own inner experience but could not find 'other than a bundle of sensations'. He did indeed find these sensory experiences to be identifiable as his own, but failed to notice that he also had an experience of ineffable ownership of them, one not identifiable with any sense experience. The ownership he was sure of, but ignored, was thus an intrinsic part of his experience, but one with "an immediate experiential givenness" as Zahavi puts it (Zahavi, 2006, p5). A similar route to Hume's was taken by a later British philosopher Gilbert Ryle (Ryle, 1950), whose pivotal book could have been titled Banishing the Ghost in the Machine. He writes, in p15 of his book The Concept of Mind: 'I shall often speak of it' [the official dogma]' with deliberate abusiveness, as 'the dogma of the Ghost in the Machine''. His claim that there is no ghost (as would correspond to the pre-reflective self) was based on linguistic analysis that is contradicted by one's own experience, again as being sure of being owner of one's experiences. The component of one's conscious experience providing such surety is subtle indeed, as we see from its minimal function and problem of interaction with the content of experience, but it is not a ghost as Ryle claimed. Similarly such an extra component of one's consciousness was missed by Brentano (Brentano 1874) who had influentially claimed that consciousness was always consciousness of, never just consciousness. But now the claim is that both sorts of consciousness exist.

We should also note that in a recent influential book (Metzinger, 2004) the author strongly claimed, as the title of the book indicates, that there is no inner self or ego. However such an egoless model is apparently contradicted by the careful and valuable dissection of consciousness developed by the author in his book; this is carefully adumbrated in (Zahavi, 2005a). Thus the No-Self of the book's title should be considered as the component of consciousness missed by David Hume, Gilbert Ryle and Brentano, being the pre-reflective self.

We should also add here that one of the major problems for a scientific basis of Western phenomenology is the strong dependence of followers of that approach on the 'bracketing' of experience, so driving away the amazing knowledge of the outside world that has presently been obtained by recent research (as emphasized recently in Gomez & Sanz, 2011). It is this process which we can, however, neglect, but preserve more specifically the transcendental nature of the phenomenological analysis of mind, providing strong support for the existence of an inner or pre-reflective self. What is missing from such a component is a way of explaining its transcendence by getting behind it by introducing unobservables. This is much as was done at the beginning stages of quantum mechanics by the introduction of the wave function. We regard the latter as equivalent to the inner self. However we will go further by introducing the CODAM model, so as to provide a more complete brain-based structure that will allow an understanding of crucial experimental observations (see sections 8, 9 & 10). It is through CODAM that we will be able to probe the inner self, equivalent to using the Schrodinger equation to probe the nature of and effects of the wave function in certain experimental situations (in spite of the unobservability of the wave function itself).

The manner in which these two components of consciousness -content and pre-reflective self- can interact is highly problematic. It was not resolved by the phenomenologists, who seemed to take the majority view that the pre-reflective self arose from the body. That is also a highly controversial position, with the danger that subjects who have lost proprioception should lose their inner self; this is known not to happen, so making that position suspect (Cole & Paillard, 1995). Also sufferers from Cotard's syndrome (where the subject claims they are dead or missing their blood or internal organs, in association with depression or suicidal tendencies) provide further contrary empirical evidence (Metzinger, 2004). Such patients have at least loss of a body image and of emotional content (related to their depression) but more extremely may no longer use I to talk about themselves: they appear to be losing their pre-reflective self. This appears to be independent of the loss of body image. This question is also discussed more generally, in terms of experimental evidence for or against the existence of a possible preflective mental state involving proprioceptive feedback, in (Taylor, 2007). However in (Damasio, 2000) evidence is presented that loss of the body image does cause a loss of phenomenal consciousness. Yet any body-based self crucially does not have the automatic property of 'immunity to error through misidentification of the first person pronoun' (Shoemaker, 1968), regarded as a crucial component of consciousness. The phrase denotes the fact that I cannot sensibly ask you 'Are you sure it is you in pain?' when you tell me you are feeling pain: you just know it is you. Thus the problem of understanding how content and owner can interact is still open, and in some ways can be seen as hard as that of explaining consciousness more generally. However it has reduced the task to something more specific in terms of the supposed two components of consciousness.

Yet the problem of understanding how the owner of conscious content becomes aware of that content is decidedly non-trivial. This is partly because the owner or I is itself not properly understood. It is possible to consider the nature of I as expounded in the writings of Eastern meditators, who have spent decades meditating on the nature of their inner self, and have built up a vast literature on their experiences across a range of disciplines, such as Zen. However this material, whilst potentially very valuable, is also solely descriptive and highly personal, thereby making it difficult to relate to modern brain science. The increasing number of brain imaging experiments performed on those who are meditating is gradually helping to bridge this gap, so that this vast literature from meditators could ultimately help exploration into the inner self. This process was also helped by the work of (Varela, 1996) who gave an initial attack on the brain basis of meditation; a further discussion is given in (Taylor, 2002a, b), where an attention-based neural model was proposed of the pre-reflective self and how it could explain the mechanism of the pure conscious experience reported by meditators.

This bipartite division of consciousness is now increasingly being accepted as an important avenue to attack mental disease, especially schizophrenia (Sass & Parnas, 2003; Taylor, 2010). The distorted mental experiences reported by patients and occurring as a major part of the disease are now being seen, through a broad study of many sufferers, as arising from an imbalance between these two components of consciousness (content and the inner self). Thereby an important brain basis for this two-component approach is being built, through increasingly detailed analysis of the accounts of schizophrenics, which should considerably strengthen and test the bipartite understanding of consciousness.

Can we develop a formal theory of the pre-reflective self? This is not trivial to achieve, and many have skirted around this task of a self able to view itself through monitoring what is occurring in its own experience. This is made more difficult by the nature of that experience, since that ends up, in the hands of Sartre, for example, as being 'nothingness' (Sartre, 1943), a feature we would accept. However that does not mean that there is actually nothing there in the content of the experience of the pre-reflective self. Such 'nothingness' can be explained as arising from brain activity at the highest coding level, but unconnected to lower level cortical activity. This lack of such activity coded at a lower level appears necessary to explain how 'nothingness' could be created; it implies that lower level neural coding in cortex gives the content to conscious experience, with its absence corresponding to no detailed experience.

It is clearly difficult, therefore, to create a formal theory of such absence as arises in the pre-reflective self without taking guidance from the brain activity underlying such 'nothingness'. We cannot easily proceed until we have a brain-based theory allowing such guidance to arise. This question will be discussed later, after we have selected such a theory.

Another influential bipartite division of consciousness was introduced in (Block, 1995). This proposed the two components of consciousness as being phenomenal consciousness and access consciousness (P and A for short respectively, as used by Block). P consciousness is just experience, granting the subject the experience of 'what it is like to be' (Nagel, 1974). Thus a P-conscious state is an experiential state. One function of P consciousness was suggested as integrating the outputs of specialized sensors and feeding the results to the executive system. On the other hand A-consciousness involves states that are poised for use as a premise in reasoning and for rational control of response.

There has been much debate of these two different bipartite approaches to consciousness, although little attempt known to the author to compare and contrast them. It is clear that they are quite different ways of slicing up experience. Thus both P and A consciousness possess content, so that these have been fused together by the Western phenomenologists. However the owner of experience posited by the Western phenomenologists does not appear in its present 'no-content' guise anywhere in the P-A split of (Block, 1995). The latter would thus seem to be vulnerable to the infinite regress of (Shoemaker, 1968). It is also to be emphasized that the owner or pre-reflective self may best be regarded as agency-free (Taylor, 2002a, b, 2009). Thus the owner cannot initiate any action, except that of standing as the guardian at the gate to consciousness, to prevent any incorrect intrusion of a distracter into consciousness and so provides the important feature of 'Immunity to error through misidentification of the first person pronoun' (Shoemaker, 1968). This phrase describes the feature raised before, that I know if I am in pain; you cannot ask me "Are you sure it is you who is in pain?"

The question of the relation of ownership to agency needs to be considered further, since the experience of agency ('did I cause the movement or did he/she?') and its possible defects have been seriously proposed as the basis of schizophrenia, for example (Frith, 1992). This appears to contradict the suggestion made earlier that such deficits arise from an imbalance in the relation between the owner and what is owned (the content). There have been several useful discussion of agency in action, such as in (Gallagher, 2000), and its possible breakdowns as might appear in schizophrenia. If however there is an attention basis for the owner/content consciousness system, then there is no agency here. This is known from discussions going back to William James. As Bricklin pointed out (Bricklin, 1999, p92) "Nothing in James' paradigm suggested that the predominance of consciousness of one thought over another was generated by the power of attention; rather the predominance in consciousness was itself attention".

This lack of agency of attention can be seen more clearly by the experimentally-based deconstruction of the attention control system into a goal-biasing system guiding the movement of the focus of attention, itself generated by a further control module, sending a signal back down into lower level posterior cortices. This bias is what causes the movement of attention and that itself is fed by values of goals, stored elsewhere. Moreover, as pointed out in (Taylor, 2002a, b) there is the added danger that if there were agentive experience in attention movement then there would be danger of an infinite regress. This would arise on exploring the source of the consciousness attributed to the attention mover.

In conclusion there is an important difference between agency ('I caused it') and ownership ('it was my thought'). Both may break down in schizophrenia, leading to a range of possible symptoms and distortions of experience (Taylor, 2010). These are thought to be explicable along the Western phenomenological division of consciousness (Sass & Parnas, 2003); a similar explanation along the P-A division is not presently available.

There are numerous other ways of slicing up consciousness, some of which will undoubtedly help in clarifying what it is that any brain-based model must explain about the complexity of consciousness. For example the popular Higher Order Thought model (Rosenthal, 1997) supposes that a thought becomes conscious by having a higher order (not necessarily conscious) thought focus on it. This also has a danger of the infinite regress when it comes to defining the self, as do other single-component approaches to the mind. It would seem presently that only by identifying a separate part of consciousness as the inner self can the infinite regress be halted.

6. Possible Brain-Based Solutions to the Mind-Body Problem

A range of brain-based solutions have been put forward to solve the Mind-Body problem. It is only through some such brain-based approach that scientific progress can be achieved, and therefore all such possible approaches are of importance to be considered most carefully. There are those who might claim that such a brain-based approach for exploring consciousness is too restrictive (Bermudez, 1998; Hutto, 2011). Yet what else can there be but the brain for creating consciousness, especially after our discussion in section 3? To claim that the brain is not relevant to a proper discussion of consciousness also misses the point: without a sure physical basis for consciousness creation, the discussion can (and has been) somewhat interminable, with many, many points of view but no one of these wins out. This is because there is no experimental basis from which various points of view can be decided between by fitting experimental data and progress can occur in a properly scientific manner.

There is only space here to briefly survey a few of these physically-based approaches to the creation of consciousness by the brain, which are summarized in Table 1.

Name of Model	General Description	Reasons Pro	Reasons Con
1) 40 Hertz Crick & Koch, 1991)	40 Hz synchronization across modules used for feature binding into object representation	40 Hz Observed in attentive object processing	1) 40 Hz also observed in anaesthetized animals 2) No reason for inner self to emerge 3) Abandoned by proposers
2) Neural Coalitions	Activation of	Gives general	1) No clear

(Crick & Koch, 2003)	suitably elaborate neural networks in the brain create consciousness	framework for consciousness	specification of an elaborate network 2) No reason for inner self to emerge
3) Recurrence (Lamme, 2003, 2006; Pollen, 2007)	Recurrent feedback causes holding of activity and hence creates consciousness	1) Recurrence crucial for brain processing (Spivey & Dale, 2004) 2) Detailed analysis possible in some cases (Wang, 2001) 3) Basic to the creation of short-term memories (see Taylor, 1999 and references therein)	1) Numerous recurrent networks observed in the brain irrelevant for consciousness creation eg hippocampus 2) No reason for inner self to emerge in this manner (through recurrence)
4) Quantum models (Penrose, 1989)	General basis of models suggested as avoiding Godel's theorem, which seems to imply that the human mind is more powerful that an sequential computational (Turing) machine	Quantum mechanics is at the basis of all dynamical system; perhaps can be extended to quantum gravitational models of the Universe (more complete), but none presently available.	1) No detailed structure for the models yet 2) No reason why the inner self is generated in this manner
5) Global Workspace (GW) (Baars, 1997)	Neural activity accesses global workspace, and thence leads to universal report across the brain, and thence to consciousness	1) Numerous sites for continued activity in higher brain centres (Duncan, 2001)	1) No place specifically made for attention, known to be crucial component for consciousness creation 2) No mechanism for inner self
6) Complexity of Activity (Edelman & Tononi, 2000)	1) Consciousness incorporates a wealth of information, only possible from complex networks 2) Can provide numerous definitions of complexity of networks	1) Increasing complexity in higher cortical sites observed 2) All definitions of complexity can be evaluated on neural networks	1) V1 is highly complex, but no consciousness of its activity 2) Difficult to measure complexity of complex networks of the brain 3) No natural inclusion of the inner self, or even of attention
7) Relational Mind (Taylor, 1999)	Mind arises by comparison of ongoing brain activity with past memories	Constant activation of relevant memories in moving around in an environment	1) No satisfactory inclusion of attention 2) No obvious emergence of an inner self.
8) Higher-Order Thought (HOT)	Consciousness arises by a second-order	Expect consciousness to be higher-order	1) Not obvious why 2nd order thought just

(Rosenthal, 1997)	thought 'thinking' about a first order one, which thereby becomes conscious	compared to lower-order sensory processing	'thinking' about a lower order one causes the latter to be conscious 2) No natural way that attention or inner self enter the process
9) Attention Copy Model (CODAM) (Taylor, 2000, 2003, 2006, 2007, 2011a, b)	1) The inner self arises from a corollary discharge of the attention movement signal 2) This 'inner experience' to be added to the coding of content as in the GW model	1) Explains ideas of Western phenomenology of Husserl et al (Zahavi, 2005b), especially of temporal flow 2) Inner self created by the corollary discharge and can be investigated, especially in schizophrenia. (Taylor, 2010)	No clarity over experimental evidence for existence of corollary discharge of attention (but see Taylor, 2011a,b)

Table 1. Brief Survey of Neural Network-based Models of Consciousness.

There are models that employ attention formally to help process information more efficiently, as in the GW model of (Baars, 1997) or in the Competing Coalitions model of (Crick & Koch, 2003), both mentioned above. However there do not seem to be any models presently available of the attention copy type, with the copy signal being used directly to provide the experience of ownership, identified with that of 'I'', other than that of CODAM, so we consider only that further, However we can see from the table that CODAM itself depends on aspects of the GW and complexity models as well as being a form of the HOT model.

8. The Control Nature of Attention

It has been proposed, as basic to the CODAM model, that there is present in association with the attention control signal a copy of that signal – the so-called corollary discharge signal – which is used to provide further control of attention movement itself. But what is it that we should be looking for as evidence of the existence of such an attention movement corollary discharge? To begin with let us consider if there should be present any corollary discharge at all. We know that the biased competition model of attention (Desimone & Duncan, 1995) is supported in general by functional magnetic resonance (MRI) analyses carried out by various groups (Bressler et al, 2008; Dosenbach et al, 2007, 2008; Corbetta et al, 2008) in humans. Already in 2001 it was reported that "Regions in the intraparietal sulcus, superior temporal sulcus and dorsal frontal regions were implicated in an attentional control circuit that may bias activity in the visual processing regions of cortex representing attended regions of space prior to the appearance of the target stimulus" (Hopfinger et al, 2001).

Such data indicates the control nature of attention, involved as it is with a feedback bias signal guided by prefrontal cortices to determine the new target for attention feedback signals sent to posterior sensory cortices. In particular there is strong experimental evidence that the sites of the control of attention movement are based in prefrontal and parietal cortex. More specifically such control is exercised from the superior parietal lobes (SPL)/intra-parietal sulcus (IPS) components in parietal lobe as well as possibly frontal eye fields (FEF) (and other regions in the prefrontal cortex (PFC) providing a goal bias for attention) (Greenberg et al, 2010 and earlier references therein; Schenkluhn et al, 2008). Support for this was also obtained by the Granger causality-based analysis of (Bressler et al, 2008). The Granger causality approach is based on delayed regression of activity of one time series onto another, and was used by (Bressler et al, 2008) to determine the most appropriate time lags between various areas as observed by fMRI. This method showed that there was very likely a flow of activity from the prefrontal and parietal areas down to various of the lower level visual cortices, as appropriate if the higher areas where in causal control of the lower ones. There was also support from results on single cell activity being amplified in lower cortical regions by an attention signal from higher cortices in monkey (Gregoriou et al, 2009). Moreover the timing of this feedback signal was observed to be after the N1/P1 (negative and positive scalp electrical waves detected 100-200 msecs after stimulus onset) period in posterior visual cortical sites involved in stimulus feature analysis. There are in fact attention feedback effects observable in amplification of the N1/P1 signal, although these appear mainly to arise from preset top-down attention control signals, and not associated with on-line modifications. It is these latter that are of main interest here as involved in the dynamics of attention movement.

We can thus begin to extend and modify the biased competition model of (Desimone & Duncan, 1995) by a more standard control model, in which it is the site of the generation of the attention movement control signal that is regarded in the brain as the crucial component for attention movement. The competitive processes ongoing in lower visual cortices are thus assumed as part of the overall (even non-attentional) input processing; here we concern ourselves with what is recognised now as the higher-order control system (Corbetta et al, 2008). As noted in (Greenberg et al, 2010, p14330) "A network of prefrontal and parietal cortical regions is thought to be the source of control signals that resolve the voluntary deployment of attention (Kastner & Ungerleider, 2000)". Furthermore from that reference "…mSPL may be the source of a domain-independent control signal that initiates the reconfiguration."

Therefore we extend here the biased competition control model of the movement of attention to that shown in figure 1.

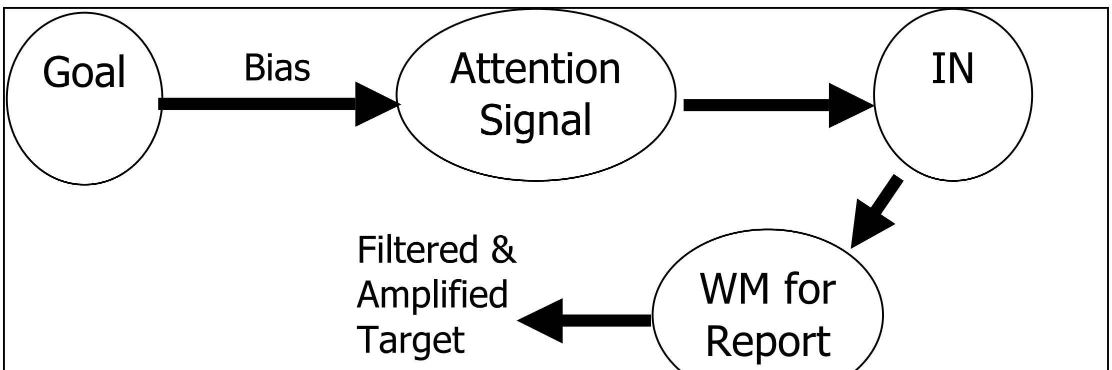

Figure 1. The Extended Biased Competition Model of Attention

In Figure 1, input, denoted IN, is modified by the attention feedback signal from the module marked 'Attention Signal' so as to amplify neural activity representing the sensory target to be amplified; this feedback signal is biased by further feedback from the module denoted 'Goal'. At the same time the top-down attention signal has been observed to inhibit neural activity representing distracters in the sensory field. The suitably filtered target neural activity is then able to access the module denoted 'WM for Report' so as to be available to other associated modules for higher level processing or for direct report as in a motor response. (We add that the Desimone and Duncan 1995 original model was more generally for attention control and not just the movement of attention; here we consider the movement of attention only, as that most relevant to the basic dynamics of attention and to the CODAM model of consciousness.)

In the model of figure 1 we have included a module, denoted 'WM for Report', acting as a visual short-term memory (VSTM), so as a short-term receptacle for the activity representing the attended stimulus. Such a further module is observed present in attended processing as the posterior component of the working memory system (Todd and Marois; 2004; Xu & Chun 2006) and in the observations on the Event related Potential (ERP) denoted as the sustained posterior contralateral positivity (SPCN - to be discussed below) associated with the access and holding of activity on the working memory site (Vogel & Machizawa. 2004; Robitaille & Jolicoeur, 2006). The Goal and the Attention Signal Generator modules are further parts of the attention control circuitry in the observed fronto-parietal network reported on in the numerous references cited above.

We can specialise the control structure of figure 1 by splitting it into several control networks, each involved in different input features, such as spatial and colour-based, following (Greenberg et al, 2010). However we do not follow that in detail here.

In addition the control model of figure 1 needs further extension to include both endogenous and exogenous attention as fused together. Indeed there are paradigms in which both forms of attention are present, so the exogenous component must be included in our discussion. This can be done using the results in (Corbetta et al, 2008) and that of earlier researchers referenced there that exogenous attention control very likely involves some form of breakthrough from the ventral route, especially involving the tempero-parietal junction (TPJ), into the dorsal attention control network (of FEF and SPL). A more complete enlarged control network can be constructed (Taylor, 2010a), but does not modify the overall strategy as indicated by the control architecture of figure 1. In particular the ultimate decider of the focus of attention is still proposed to be the SPL/IPS complex, with bias from the prefrontal cortex, as well as from suitable saliency-type maps in the exogenous case.

Having justified from an experimental point of view the basic control model of figure 1, we conjectured the existence of a corollary discharge component of this basic attention movement control signal (just mentioned) to be sent to other sites in cortex at about the same time or just a little later as the signal itself is sent to posterior cortices (Taylor, 2000). Support for the existence of such a signal for attention movement is to be obtained in a similar manner to that in motor control, for which the evidence for such an efference copy was strong (Desmurget & Grafton, 2000; Crapse & Wurtz, 2008).

There is a clear distinction of such a corollary discharge signal from the feedback attention control signal, which has as its function solely to amplify lower level activity associated with the activity of the attended stimulus and inhibition of nearby distracters. Such a backward-going attention control signal has been observed at multi-unit activity level, for example in

(Mehta et al, 2000) and in many ERP paradigms. We now extend this model to that of figure 2 by assuming that a corollary discharge signal is also generated (and observable) in attention.

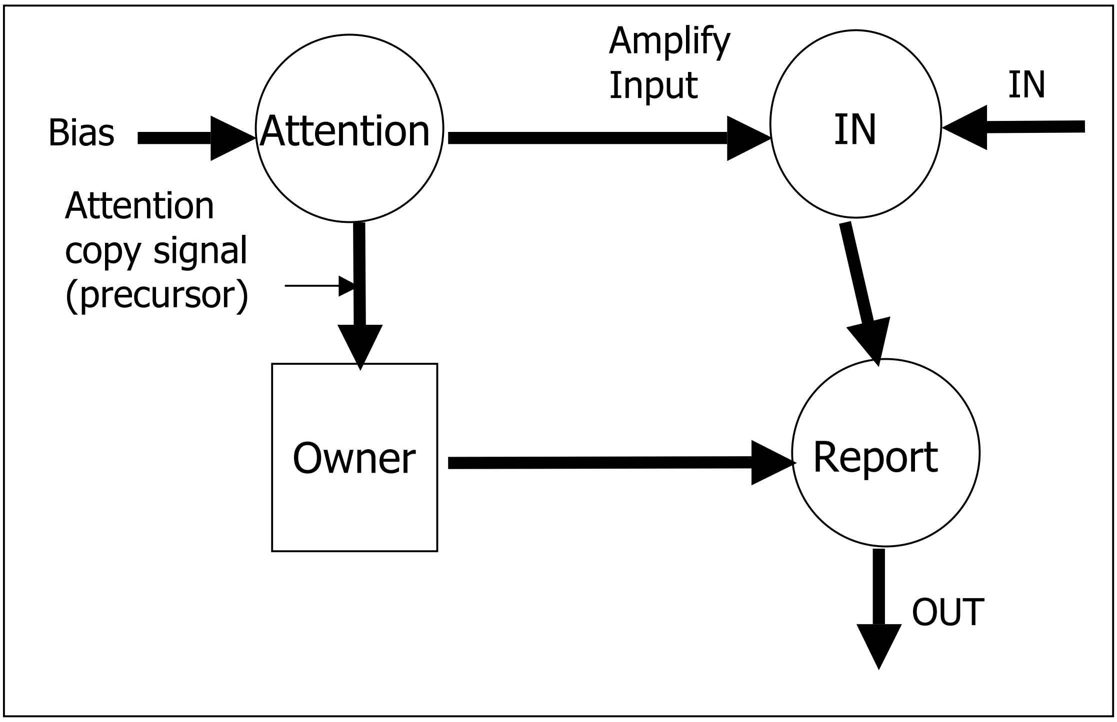

Figure 2 Extended Attention Control Model with Corollary Discharge Signal

The corollary discharge signal is denoted as 'Attention copy signal' in figure 2. The additional module denoted 'Owner' is a short-term memory site on which to hold the corollary discharge so as to be able to use it during the period when further manipulations may need to be performed on activity in the working memory module. The owner module can be regarded as a predictor of future activity about to arrive on the working memory site, and is well used in engineering control models in many applications.

In figure 2 the attention copy signal (the term corollary discharge implies a transformation of the actual attention signal itself, so beyond a copy) was proposed (Taylor, 2000) as being emitted by the attention movement controller (in the superior parietal lobe (SPL)/infra-parietal sulcus (IPS)) and sent to a short-term buffer for use in: a) Inhibiting distracters on the buffer; b) Amplifying, so speeding up, the arrival of activity from the attended target stimulus onto the buffer, acting as a 'Report' module. The short-term corollary discharge buffer in figure 2 is termed the 'Owner' module, since we will later relate it to the inner self or ownership component of conscious experience.

It has already been pointed out in several publications on CODAM that the temporal flow of activity in the model supports and explains that noted by Husserl and many others on the Western phenomenological side (Zahavi, 2005b; Sokolowski, 2000). In particular the 3 temporal stages of protention, the primal impression and retention would be created by activity in the Owner module (as a predictor, so corresponding to protention), the relevant short-term memory buffer (on gaining access to the site for report, so being the primal impression) and continued, but decaying such activity (so being the retention component (Sokolowski, 2000). The temporal order of these components agrees with that expected by the Western phenomenolgists.

9. Towards a Formal Theory of the Pre-Reflective Self

We can now return to the question raised earlier in section 6: can we develop a formal theory of the pre-reflective self. We can now employ CODAM as a brain basis for the flow of activity during the dynamical flow of brain activity leading to, during the experience of, and following, consciousness. From CODAM the pre-reflective self arises from brain activity just before consciousness, and associated with the detailed monitoring of activity on the working memory site.

There are two levels at which can proceed: one is at the purely neural activity level, the other at the experiential level. The former of these may be modelled, by CODAM, as a control system for attention; the equations are given in detail in (Fragopanagos et al, 2005), where a simulation is performed of the CODAM model for the attentional blink. We refer the interested reader to the relevant pages.

The second of the levels above, that of experience can now be attempted. The first step must involve the formal structures of experience involved in attention processing; we already developed a two-stage process in going from figure 1 to figure 2 in the previous section. Let us consider an environment E containing objects S(i), where i runs over the set of objects observable by the subject. Let S(a) denote the particular object in E which will be made conscious to the subject by a control model of attention, such as that of CODAM. Each of the objects S(i) will cause brain activity we denote by CS(i), C standing for the processing leading to the code for S(i) in the brain. Then CS(i) will be represented by neural activity representing the features of the various stimuli, being distributed across the hierarchy of occipital lobe modules for vision and similar (but much reduced) codes for auditory and other senses. Further let Š denote the set of all object codes CS(i):

$$\check{S} = \{CS(i)\} \qquad (1)$$

Singling out the attended object is the process of attention in the subject's brain. Then the map

$$\check{S} \rightarrow CS(a) \qquad (2)$$

is the basic attention map which causes the brain activities representing the object S(a) to be selected as attended to. It is assumed that this process leads to activity on the working memory site to allow report to be spread about to the other working memory sites, as in the GW theory of (Baars (1998), and as briefly described in the previous section 7. We denote the working memory site brain activity arising in this manner from the external object S(i) by WMCS(i).

Such a formal theory of attention is not yet one of consciousness. For although there is report of S(a) around the appropriate working memory centres in the brain, there is no one who is conscious of this activity, so no owner. Following CODAM, we define a monitoring process acting on Š so as to destroy distracter stimuli entering attention. Thus this monitoring process must depend on other neural activities occurring at a subliminal level in low level cortices to single out the stimulus code CS(a) as activating the relevant working memory codes for those inputs, as

$$\check{S} \rightarrow Mon\{\check{S}; a\} \qquad (3)$$

Such a monitoring process as given by Mon is complex, depending, as it must, on those distracters which have entered the working memory site concomitantly with the activity coding for the stimulus S(a). This set of stimuli on the working memory site need not be very large, since filtering of many unwanted distracters by attention will already have been achieved at a lower level through inhibition from the attention control signal to the lower level cortical sites.

In all, the formal structure of the attention process leading to consciousness, with an owner, has the temporal structure

$$\check{S} \rightarrow \mathrm{Mon}\{\check{S};\, a\} \rightarrow \{S(a){:}\ LLS(a)\} \rightarrow WMCS(a) \qquad (4)$$

The process indicated by the first arrow in (4) is that of removal of any remaining distracters from the working memory site (and possibly other processes). Naturally it depends on all of the objects in E (as possible distracters, as well as the attended object codes). However it has no relationship to lower level coding of any of these objects. It is for that reason that it can be considered as a 'nothingness', since it is coded at such a high level that there can be no detailed content present in it, nor any experience of any content. Thus the activity represented by Mon (Š, a) can represent the pre-reflective self. It is clearly complex dynamic (changing in time) activity, is related most specifically to the attended object S(a), but has no related low-level activity to give the activity any conscious content. However being on a working memory site it has the character of being able to term it 'conscious', although in a rather thin manner. As discussed elsewhere (Taylor, 2010) there may be activities from other high-level sites (such as reward attributions in the orbito-frontal cortex) which are also feeding into the working memory site, so giving the owner an emotional slant in the reward case.

In the process given by the transition of the first arrow of (4), the pre-reflective self is completely determined by its monitoring transformation properties it is carrying out, and depends importantly on the attended object S(a). It can thereby be regarded as the owner of the neural activity of the attended object, and hence, in experience, the owner of the conscious experience itself.

The second arrow indicates a second stage, when the owner had done its work on the working memory site and leads to a unique activity. This latter has content by virtue of its interconnectivity with lower level codes filling out the cortical activity on the working memory site. This activity is denoted by LLS(a) in (4) (LL denoting 'lower level'). This connectivity may be achieved in a variety of ways, such as by the use of 40 Hz synchronisation (Crick & Koch, 1991) or by more direct attention control. This activity is expected to stand out (by amplification or synchronisation) in lower level cortices from that for distracters. The final arrow denotes the holding of the neural activity, coding for the attended object, on the working memory site.

We claim that (4) is a brief formal expression for the development of consciousness of the object S(a). It occurs in time, so that the first arrow may take some hundreds of milliseconds before it is completed. It is during that period that the pre-reflective self, the owner, is present, as just argued. Only when that stage has been achieved will the next stage, specified by the second arrow in (4), be able to proceed. Moreover the three stages in (4) can be identified with those of protention, primal impression and retention of Husserl (as already noted). We note in (4) the removal from conscious content of the working memory activity of S(a) at the final stage, since there is no longer the lower level activity expected to be coupled to the working memory trace activity. However we expect that can be re-activated at its original level by suitable rehearsal mechanisms (Korsten et al, 2006). Thus equation (4) is a formal expression for the overall creation of consciousness as evinced by the studies of the school of Western phenomenology (Sokolowski, 2000; Zahavi, 2005).

The above discussion may be objected to as falling between two streams, one being that of detailing the flow of neural activity under attention control, the other being that of the conscious experience of the subject. More precisely what we have done is to give a conscious 'gloss' on the neural activity represented by (4). It might be argued that such a method is fraudulent in trying to bridge the mind-body gap. However we claim we have done more than that: we have provided a dictionary to translate between the two levels of description. One, the neural level, is specified by the process of (4). The other is that of the subjective and phenomenological experience of the subject. It is given by the descriptions of their experience by subjects built up from many experiments in brain science, psychology and the philosophy of mind. In particular it dovetails nicely with the results obtained by the philosophers of the school of Western phenomenology, as already noted. With (4) and the dictionary we thus have achieved our aim of setting out an initial framework for a formal analysis of the pre-reflective self.

10. Experimental Support for Attention Corollary Discharge

This section involves many acronyms, for which the author apologises. However these are part of the scientific literature. It is difficult to write succinctly about these brain potentials without resorting to the acronyms without much turgid and lengthy discourse. However their impact will be reduced as much as possible by spelling out in full the acronymic descriptor when first they appear, so as to ease the readers' passage through this section.

10.1. Criteria for a Corollary Discharge

Various properties of a corollary discharge signal must be specified before we can try to discover if such a signal has already been observed in experimental data. In other words we need roughly to know what we are looking for before we set out on our quest for a corollary discharge of attention movement in the brain, as was mentioned above. We take some guidance from the successful application of such copy ideas to motor control models (Wolpert & Gharharmani, 2000; Desmurget & Grafton, 2000). We can also expect the existence of a corollary discharge as supported by the pre-motor theory of attention (Rizolatti et al, 1987). That approach considers attention movement as possessing considerable similarity to that of motor action, although now restricted to possibly different regions of the brain, with some considerable overlap being presumed of the two sets of functional areas.

We use the criteria suggested by (Wurtz & Sommer, 2004) in their search for the corollary discharge of the motor control signal (movement of the eyes accompanied by a signal from the superior colliculus to the thalamus and thence to the frontal eye fields). We must however modify the criteria C1-C4 of (Wurtz & Sommer, 2004) as follows:

A1) Relevant signals originate in an attention control area;

A2) These signals precede and spatially represent the attention movement;

A3) Eliminating these signals does not impair attention movement in tasks not requiring corollary discharge;

A4) Eliminating the signals does impair attention movements in tasks requiring corollary discharge.

The criteria A1 and A2 can be applied directly to any observed attention movement control process. A3 and A4 require considerable care in that the elimination process itself may need to be done on humans (without damage to their brains). This can be achieved in one way by means of trans-cranial magnetic stimulation (TMS), as has been achieved in a number of relevant experiments. An alternative source of data is that of patients with brain defects, especially in stroke: such patients are well-known to have defects in the attention system, as many researchers have argued.

We must add that there is still uncertainty about the nature of the proposed attention copy signal. Is there just one such signal to be searched for or are there different signals being involved in different functions associated with attention (but not with direct modification of lower level sensory processing)? It is feasible that there may be a number of different functions carried out by such a copy: amplification of the target representation as by refreshing its working memory representation, inhibition of possible distracters entering the working memory, activating reward values for the target, being used for error correction given an incorrect target. There may also be other uses. This is consistent with the increasing set of components of higher-order information processing now being uncovered in the brain.

10.2. Looking for the Relevant Signal(s)

We must now turn to the set of signals in the brain we may use to provide evidence for the existence of the corollary discharge of attention movement control signal. To do that we have first to search for the actual, attention movement control signal itself, and determine its timing. That signal would be one sent to the lower posterior sensory cortices so as to amplify the neural activity representing the relevant target stimulus, with associated inhibition of any distracter, especially if they are near to the target. That will provide us with a time for any possible beginning of a corollary discharge signal.

There seems to be an apparent contradiction, however, between functional magnetic resonance imaging (fMRI)/positron emission topography (PET) and electrical encephalography (EEG)/magneto encephalography (MEG) signals as to when exactly the focus of attention can be observed to move. For example there is the set of components observed under certain paradigms described in (Stigchel et al, 2006; Kiss et al, 2008), with acronyms such as EDAN, ADAN, LDAP, RLIP, or those event related potential (ERP) signals detected by other paradigms, such as the distracter positivity P(D) and the target negativity N(T), in (Hickey et al, 2009) and suggested to be components of the important posterior contra-lateral brain activity N2pc (considered as an indicator of a shift of attention), or the later positively-sited contralateral temporal lobe activity, denoted Ptc observed in the paradigms of (Hilmire et al, 2009), as well as the even later sustained posterior contralateral negativity, denoted SPCN, of (Robitaille and Jolicoeur, 2006) and also observed by other research groups. Which of these signals (or possibly more than one) could represent an attention movement corollary discharge signal?

From the above plethora of ERP signals, we cannot conclude there is one and only one such corollary discharge signal to be searched for. But we expect to have to look outside primary sensory regions and after the initial attention movement control signal has been sent to sensory cortices, for any signal that can be suggested as a corollary discharge of this primary attention movement control signal. The results of (Kiss et al, 2008) would seem to indicate that only the anterior-directing attention negativity (ADAN) and late-directing attention positivity (LDAP) are good candidates for such a signal: these are the only signals they observed in their paradigm in which either an informative or a non-informative cue appeared before a target. In the former case a clear ADAN signal was seen, whilst for the latter this had vanished. Moreover when a target appeared a very clear N2pc was observed for both the informative and non-informative cue cases. This seemed to indicate, as they claimed (Kiss et al, 2008, p247) that attention had already moved to the necessary hemisphere before the N2pc, and that "…the N2pc is instead linked to processes that occur after such shifts have been completed." However (Eimer, private communication, 2010) has noted that "…we are still looking for this elusive signal (of the movement of attention)" implying that more work needs to be done to pin it down.

This difficulty arose because the relevant LDAP posterior cortical ERP signal, in the Kiss et al case (Kiss et al, 2008, p244) seemed to rule out the N2pc (since the N2pc was the same in both the informative and the uninformative cue cases) but furthermore "was absent with uninformative cues [] but only revealed a trend towards the presence of the LDAP in the informative cue condition that fell short of statistical significance…" There are numerous sightings of the LDAP by other research teams, but none in association with the crucial informative/uninformative cue comparison.

10.3. Experimental Support for the Existence of an Attention Movement Signal from MEG Data

We consider the time line of activity to be expected from the extended attention control model of figure 3, as follows:

0 (stimulus onset) → 50-200 msecs (feature analysis, possibly attention modulated by a previous attention control signal, and further modulation by local low-level saliency processing) → 180-200 msecs (attention control signal generation) → 220-350 msecs (corollary discharge, before report occurs, with some removal of distracters) → 350-1000 msecs (attended stimulus access to its short-term or working memory, to be available for report, and possibly involving slightly earlier further corollary discharge activity enabling more complex distracter removal).

There are several sets of relevant data known to us to explore the time line (1), and in particular the signal for the movement of attention. The first of these was obtained a decade ago (Hopf et al, 2000) by MEG for subjects performing a target search task that produced a difference signal between the two hemispheres. In particular there was a difference in the activity contralateral to the target as compared to that in the same hemisphere.

The paradigm consisted of two sets of twelve short horizontal and vertical lines, one set on each side of a central fixation cross. Subjects were asked to fixate on the central cross and attend to either a red or a green line (the other 22 lines being blue distracter bars, and the red and green lines being in opposite hemispheres) and indicate whether the attended line was horizontal or vertical in each stimulus array (with the colour to be attended being specified at the beginning of each block of 200 individual trials, with 12 trial blocks in all). The stimuli for a given trial were on for 750 msecs, and there was a variable-duration gap of 600-900 msecs between each stimulus presentation for response. Stimulus confounds were avoided by using red and green lines in each of the stimulus trials.

The resulting difference signal between the two hemispheres was observed to have two components: one an early one at 180-200 msecs in SPL and the second at 220-240 msecs, which was observed in the temporal lobe. We take the first of these signals to show the initial attention movement control signal being generated in the SPL/IPS complex, as supported by fMRI data (Corbetta et al, 2008; Bressler et al, 2008). This timing data was used in the time line (1). The later signal at 220-240 msecs represents the arrival at a temporal lobe working memory site, which we propose to be for a corollary discharge created from the earlier SPL signal.

One supposed problem with this data is that the early signal has not been observed in later MEG experiments performed by Hopf and his colleagues (Hopf, et al, 2002a; ibid 2004). However this discrepancy has been handled by the Hopf group by noting that the experiments reported in the papers of 2002a and 2004 very likely had 'silent' MEG signals compared to those arising in the 2000 paper. As noted in the 2002a paper

> *"…the pattern of results across experiments is consistent with the proposal that the parietal subcomponent reflects the mechanism that initiates shifts of attention, whereas the occipito-temporal subcomponent reflects the mechanisms that implement the selection of relevant information once attention has been shifted." (Hopf et al, 2002a, p27).*

We still need somehow to reconcile the (Hopf et al, 2000) claim that the early component of the N2pc is that of the signal of attention movement with the data we noted above of (Kiss et al, 2008); this latter was claimed to show that the N2pc occurs after any attention movement has occurred. The paradigm of (Kiss et al, 2008) used a circular arrangement of 12 small squares, ordered as if round the circumference of a clock, with a single one being oriented as a diamond, and with one side of the diamond cut short. The earlier cue consisted of an arrow, pointing either to the 2, 3 or 4 positions on the clock or to those at 8, 9, 10 on the other side. In the uninformative cue case there were two such arrows, pointing in opposite directions (so indeed being uninformative).

We summarise the points pro and con for the signals (ADAN/LDAP pair and the parietal component of the MEG N2pc signal) to represent the attention movement signal in Table 2 below. We have mentioned the ADAN signal briefly as a frontal precursor of the posterior LDAP signal; as such ADAN may feed into LDAP, although no causal flow from the former to the latter is in evidence.

Signal Considered	Pro	Con
N2pc (Hopf et al, 2000)	Clearly observed, and absent in later paradigms (Hopf et al, 2002a, b, 2004) as explained in those papers	Suspected by some other researchers (possibly poor source localisation modelling)
	Suitably early (180 – 200 msecs) and sited (Parietal)	Present in both the informative and uninformative cases by (Kiss et al, 2008)
	Experimentally observed as vulnerable to TMS applied to SPL (Schenkuhn et al, 2008)	
ADAN (frontal)/LDAP (posterior) Cortical ERP	ADAN observed as significant in the informative	LDAP not observed as significant in either

Signals (Kiss et al, 2008).	cue case, and not in the uninformative case.	informative or uninformative cue cases.
		Only hinted at after a double subtraction

Table 2. Which/When/Where is the Attention Movement Control Signal?

Pros and Cons for the possibility of the ADAN/LDAP pair or the N2pc signal being the attention movement control signal, as required to exist in any control model of attention.

We finally assume that the attention movement signal was generated at or before the N2pc signal at around 200msecs post-stimulus. Such timing has been proposed by the majority of investigators of the N2pc, including Kiss et al, in their 2008 paper. Thus we will search for the corollary discharge signal as created during or after the N2pc.

10.4. Relevant ERP Componens

We tabulate below possible ERP components which could play a role in the corollary discharge of the movement of attention: these all are observed during or after the N2pc, as we concluded was necessary from the evidence in sub-section 9.3.

Name of ERP	Timing	General Features and Function	Position in the cortex	Pro	Con
1) N2pcb	220-280 ms post-stimulus	2nd component observed after the earlier N2pca at 180-220 ms (Hopf et al, 2000)	Temporal lobe (as in a working memory site)	Arises after the putative attention movement control signal, N2pca, as required	May be the original copy of the attention control signal (see discussion in sub-section 9.3 concerning N2pca)
2) RLIP (right-lateralised inhibition positivity) (Van der Stigchel et al, 2006)	220-330ms post-stimulus, contralateral to cued direction	Activation in response to top-down inhibition. Counterpart of EDAN, but involves inhibitory processes evoked by the distracter cue	Maximum at P04	Used to inhibit distracters, as expected from a corollary discharge signal	Possible component of the attention corollary discharge signal
3) EDAN (early directing attention	200-400 ms post-stimulus	Oriented along cue direction for target	Posterior electrode sites (PO7/PO8)	Unclear status; needs further experimentation	For alerting to upcoming target, so could be

negativity)		preparation			main attention control signal for such alerting
4)Ptc (Hilimire et al, 2010)	290-340 ms positivity contralateral to attended item	Bias of processing to individuate task-relevant stimulus	More at temporal electrodes	Unclear status; needs further experimentation	Unclear and to be explored
5) ADAN (anterior directing attention negativity)	350-650 ms post-stimulus	Contribution by frontal brain structures to control of spatial attention shifts	Lateral fronto-central electrodes (FC5/FC6)	Undoubtedly part of corollary discharge, but unclear concerning ownership nature	Unclear, and to be explored
6) LDAN (late attention directing negativity) (Van der Stigchel et al, 2006)	400-500 ms post-stimulus	Widespread negativity for attention orienting and saccade programming prior to target presentation	Frontal and central electrodes	Unclear, especially in relation to saccade preparation (so not directly related to target being attended)	Unclear and to be explored
7) LDAP (late directing attention positivity) (Kiss et al, 2008)	Final 200 ms of the cue-target interval		Posterior electrode sites (PO7/8)	Unclear status	Unclear and to be explored
8) SPCN (sustained posterior contralateral negativity) (Robitaille & Jolicoeur, 2006)	300-900 ms post-stimulus	Clearly partially involved in distracter removal in same hemisphere, as well as late component possibly involved in motor action preparation	Posterior electrode sites (PO7/PO8)	Strong candidate for corollary discharge	Needs more analysis as to precise mechanism (Taylor, 2012)

Table 3: Summary of ERP Activities Relevant to Attention Movement Corollary Discharge Existence

Our conclusion from the above table is that there are several strong candidate ERPs to arise from corollary discharges of attention movement control signals, including the N2pcb of

(Hopf et al, 2000), but especially the SPCN of (Robitaille and Jolicoeur, 2006). Other ERP signals also have indications of involving inhibition of distracters as part of their involvement in attention movement control. However they need further careful investigation.

11. Ownership and How to Probe It Further

There are numerous important features from the above models of consciousness that we must take forward to build a more complete model with some chance of success: attention as a filter and possible source of feature binding, activity coming into awareness when accessing sites of working memory, associated attention copy ownership, various forms of memory as crucial components to give content to awareness, synchronization as a possible additional source of feature binding for object representations, suitably complex interacting brain states, emotion values as giving the whole system value and source of drive, and so on.

Ownership has been emphasized in various publications (Taylor, 2002a,b, 2007, 2011a). What is ownership and how can we probe it further? It has been already pointed out in various places noted above by the author that ownership is a subtle aspect because it involves neural activity which is coded at a high level, in particular to remove distracters from the focus of attention, and has no or very little connection with the lower level content codes of sensory cortices in the brain. It is this divorce from content which leads to the 'nothingness' claim of Sartre (Sartre, 1943) and the difficulty to detect the ownership experience itself. But if there is an experience of ownership (as claimed by the phenomenological school of Husserl et al (Zahavi, 2005b) then what is being owned and who is doing the owning? These are questions deserving a crisp answer before we can embark on further search for that owner.

The answer to the question as to what is being owned is that it is the content of the experience that is about to arise into reportability across short-term memory sites, as in the GW model of (Baars, 1997). This content is itself being worked on, as in the N2pc and the SPCN discussed in the previous section. Various distracters are being removed from lower level processing or from the VSTM (in the case of vision) before report can be achieved. It is this cleaning–up of the results of earlier processing, carried out before content awareness arises, which is what ownership appears, from this data, to be about. The cleansing process itself grants to the subject the experience of preparing the content to allowed to be reported, thus casting the stamp of ownership on it.

Besides the purely 'academic' cleaning process there may also be goal bias signals arising from emotional/reward sites (say in orbito-frontal cortex) or from competitive discriminative support from cingulate regions. At the same time memory processes or prefrontal regions may be activated to bring about rehearsal of earlier material, as needed by rehearsal to guide further actions. Thus a number of other regions could contribute to the ownership feeling. However we suggest that the primary ownership content may well arise from the basic activity of distracter removal. Indeed that is the purpose of attention itself: to allow for filtering out distracters and concentration on the brain representation of target activity in order to be able to handle further processing of only the latter. That is all that attention consists of: target search and related inhibitory processes to remove distracters.

Do these various components of brain activity begin to look as if they could provide a basis for a suitably scientific theory of consciousness? This can only be achieved by a modeling-cum-experimental attack on the many paradigms and associated observations in which consciousness crucially enters, and based on models including the three main principles mentioned earlier and some or all of the features of the various models outlined above. Such phenomena include, for example, the attentional blink, the Stroop test, attention search and associated phenomena such as the Posner benefit paradigm, processing of subliminal inputs, the phenomenon of blindsight and numerous others. Only when all of these have been explained satisfactorily at a quantitative level by such a developed model can we begin to be able to justify the claimed final solution to the mind-body problem. Even then there will be enormous problems of interpretation.

Acknowledgements

The author would like to thank his younger colleagues who have worked with him on various parts of the overall program developed here: Nikos Fragopanagos, Neill Taylor, Matthew Hartley, Nienke Korsten, and also thank Michael Woods, who has helped illuminate some of the murkier parts of CODAM from a psychological and philosophical viewpoint. He would also like to thank the EU for financial support during the completion of this work through the DARWIN (manipulable autonomous robot) project.

References

Baars, B (1997) Some essential differences between consciousness and attention, perception, and working memory Consciousness and Cognition 6(2-3):363-71

Bermudez JL (1998) The Paradox of Self-Consciousness Cambridge MA: MIT Press

Block NJ (1995) On a Confusion About the Function of Consciousness. Behavioural and Brain Sciences 18:227-247

Brentano F (1874) Psycholgie vom empirischen Standpunkt. Hamburg: Felix Meiner; English translation: LL McAlister Psychology from an Empirical Standpoint. London: Routledge, 1973

Bressler S L,Tang C M, Sylvester G M, Shulman G &.Corbetta M (2008) Top-Down Control of Human Visual Cortex by Frontal and Parietal Cortex in Anticipatory Visual Spatial Attention J Neuroscience 28(40):10056-10061

Bricklin (1999) A variety of religious experience: William James and the non-reality of free will. in The Volitional Brain, ed B Libet, A Freeman & K Sutherland. Thorverton, Devon: Imprint Academic

Cole J & Paillard J (1995) Living Without Touch and Peripheral Information about Body Position and Movement: Studies with Deafferented Patients, pp 245-266 in The Body and the Self. JL Bermudez, AJ Marcel & N Eilan (eds). Cambridge MA: MIT Press

Corbetta M, Patel G & Shulman G (2008) The Reorienting System of the Human Brain: From Environment to Theory of Mind. Neuron 58:306-324

Crick FHC & Koch C (1991) Towards a neurobiological theory of consciousness. seminars in neuroscience 2:237-249

Crick FHC & Koch C (2003) A framework for consciousness. Nature Neuroscience6(2):119-126

Damasio AR (2000) The Feeling of What Happens: body, emotion and the making of consciousness. Vintage Press.

Desimone R & Duncan J (1995) Neural mechanisms of selective visual attention Ann. Rev. Neurosci., 18:193-222

Desmurget M & Grafton S (2000) Forward modelling allows feedback control for fast reaching movements. Trends in Cognitive Sciences 4:423-431

Dosenbach N U F, Fair D A, Cohen A L, Schlaggar B L & Petersen SE (2008) A dual-networks architecture of top-down control Trends in Cognitive Sciences 12(3):99-105

Duncan J (2001) An Adaptive Coding Model of Neural Function in Prefrontal Cortex. Nature Reviews Neuroscience 2:820-829

Edelman G & Tononi G (2000) A Universe of Consciousness: How Matter Becomes Imagination. NY: Basic Books

Elliott R & Dolan RJ (1998) Neural Response during Preference and Memory Judgements for Subliminally Presented Stimuli: A Functional Neuroimaging Study. Journal of Neuroscience 18(12):4697-4704

Flanagan O (1992) Consciousness Reconsidered. Cambridge MA: MIT Press

Fragopanagos N, Kockelkoren S & Taylor JG (2005) A neurodynamic model of the attentional blink Cognitive Brain Research 24:568-586

Frith CD (1992) The Cognitive Neuropsychology of Schizophrenia. Englewood Cliffs NJ: Erlbaum

Gallagher S (2000) Philosophical conceptions of the self: Implications for cognitive science. Trends in Cognitive Sciences 4(1): 14-21.

Gomez J & Sanz R (2011) The Unbearable Heaviness of Being, in Phenomenological AI Journal of Mind Theory 0(2):175-186

Greenberg A S, Esterman M, Wilson D, Serences J T & Yantis S (2010) Control of Spatial and Feature-Based Attention in Frontoparietal Cortex Journal of Neuroscience 30(43):14330-14339

Gregoriou G G, Gotts S J, Zhou H & Desimone R (2009) High-Frequency, Long-Range Coupling Between Prefrontal and Visual Cortex During Attention Science 34:1207-1210

Haynes J-D, Driver J & Rees G (2005) Visbility Reflects Dynamic Changes of Effective Connectivity between V1 and Fusiform Cortex. Neuron 46:811-821

Hickey C, DiLollo V & McDonald J J (2009) Electrophysiological indices of target and distractor processing in visual search J Cog Neuroscience 21(4):760-775

Hilimire M R, Mounts J R W, Parks N A & Corballis P M (2010) Event Related Potentials Differentiate Effects of Salience and Space in Biased Competition for Visual Representation PLoS One 5(9):e12677. doi:10.1371/journal.pone.0012677

Hopf J-M, Luck SJ, Girelloi M, Tillman H, Mangun GR, Scheich H & Heinze H-J (2000) Neural Sources of Focussed Attention in Visual Search Cerebral Cortex 10:1233-1241

Hopf J-M, Boelmans K, Schoenfeld A M, Heinze H-J & Luck S J (2002a) How does attention attenuate target-distractor interference in vision? Evidence from magnetoencephalography recordings Cognitive Brain Research 15:17-29

Hopf J-M, Vogel E K, Woodman G, Heinze H-J & Luck S J (2002b) Localizing discrimination processes in time and space J Neurophysiol 88:2088-2095

Hopfinger J B, Woldorff M G, Fletcher E M & Mangun G R (2001) Dissociating top-down attentional control from selective perception and action Neuropsychologia 39:1277-1291

Hume D (1888) A Treatise on Human Nature Oxford: Clarendon Press

Kastner S & Ungerleider L G (2000) Mechanisms of visual attention in the human cortex Annual Reviews of Neuroscience 23:315-341

Hutto D (2011) Still a Cause for Concern: Reasons, Causes and Explanations Hutto web-site

Kiss M, JVan Velzen J & Eimer M (2008) The N2pc component and its links to attention shifts and spatially selective visual processing Psychophysiology 45:240-249

Korsten N, Fragopanagos N, Hartley M, Taylor N & Taylor JG (2006) Attention as a Controller Neural Networks 19(9):1408-1421

Lamme V (2003) Why visual attention and awareness are different. Trends in Cognitive Sciences 7:12-18

Lamme V (2006) Towards a true neural stance on consciousness. Trends in Cognitive Sciences 10(11):494-501

Levine J (1983) Materialism and qualia: The explanatory gap. Pacific Philosophy Quarterly 64:354-361

Luck S J & Hillyard S A (!994a) Spatial filtering during visual search. Evidence from human electrophysiology. J Experimental Psychology: Human Perception and Performance 20:1000-1014

Luck S J & Hillyard S A (1994b) Electrophysiological correlates of feature analysis during visual search. Psychophysiology 31:291-308

Mehta A D, Ulbert I & Schroeder C E (2000) Intermodal Attentional Selection in Monkeys I: Distribution and Timing of Effects across Visual Areas. Cerebral Cortex 10:343 358

Merleau-Ponty M (1945) Phenomenologie de la perception. Paris: Editions Gallimard; English tranlsation: C Smith. Phenomenology of Perception. London: Routledge and Kegan Paul, 1962

Moutoussis K & Zeki S (2002) The relationship between cortical activation and perception investigated with invisible stimuli. Proc Nat Acad Sci USA 99:9527-32

Munneke J, Van der Stigchel S & Theeuwes J (2008) Cueing the location of a distractor: An inhibitory mechanism of spatial attention? Acta Psychologica 129:101–107

Nagel T (1974) "What is it like to be a bat?" The Philosophical Review 83:434-50

Penrose R (1989) The Emperior's New Mind. Oxford: Oxford University Press

Pollen D (2007) Fundamental Requirements for Primary Visual Perception. Cerebral Cortex December 17, 2007

Rahnev, D. et al. (2009) Inattention boosts subjective visibility:implications for inattentional and change blindness. Journal of Vision. 9, 157

Rizzolatti G, Riggio L, Dascola I & Umilta C (1987) Reorganization of attention across the horizontal and vertical meridians – evidence in favour of a premotor theory of attention. Neuropsychology 25:33-40

Robitaille N & Jolicoeur P (2006) Fundamental Properties of the N2pc as an Index of Spatial Attention: Effects of Masking Canadian Journal of Experimental Psychology 60(2):101-111

Rosenthal DM (1997) A Theory of Consciousness. in N Block, O Flanagan & G Guzeldere (eds) The Nature of Consciousness, pp 729-753. Cambridge MA: MIT Press

Ryle G (1950) The Concept of Mind. Oxford: Oxford UNiversity Press

Sartre J-P (1943) L'Etre et le neant. Paris: Gallimard; Engilsh translation: HE Barnes, Being and Nothingenss, New York: Philosphical Library, 1956

Sass L & Parnas J (2003) Schizophrenia, Consciousness and the Self. Schizophrenia Bulletin 29(3):427-444

Schenkluhn B, Ruff C C, Reinen K & Chambers CD (2008) Parietal Stimulation Decouples Spatial and Feature-Based Attention Journal of NeuroScience 28(44):11106-11110

Sergent C, Baillet S & Dehaene S (2005) Timing of the brain events underlying access to consciosuness during the attentional blink. Nature Neuroscience Sep 11, 2005

Shoemaker S (1968) Self-reference and self-awareness. The Journal of Philosophy 65:555-567

Sokolowski R (2000) Introduction to Phenomenology Cambridge UK: Cambridge University Press

Spivey MJ & Dale R (2004) On the Continuity of Mind: Toward a Dynamical Account of Cognition. pp 87-142 in The Psychology of Learning and Motivation, vol 45

Taylor J G (2012) The Corollary Discharge of Attention Exists? Consciousness & Cognition, Volume 21, Issue 1, Pp. 325–339

Taylor J G (2011a) The Problem of I, Journal of Consciousness Studies (in press)

Taylor J G (2010). A neural model of the loss of self in schizophrenia Schizophrenia Bulletin Volume 37, Issue 6 Pp. 1229-1247

Taylor J G (2010a) The Perception-Conceptualisation-Knowledge Representation-Reasoning Representation-Action cycle: The View from the Brain Chapter 8, pp 243-285 in V Cutsuridis et al (eds) Perception-Action Cycle: Model, Architectures and Hardware London: Springer

Taylor JG (2007) On the neurodynamics of the creation of consciousness. Cognitive Neurodynamics 1:97-118

Taylor JG (2006) The Mind: A User's Manual. London: Wiley & Sons

Taylor JG (2003) Paying Attention to Consciousness. Progress in Neurobiology 71:305-33

Taylor JG (2080) Science and the Supernatural London: Maurice Templeman-Smith

Taylor JG & Korsten N (2010) Connectionist models of emotion in Oxford Companion to the Affective Sciences Oxford: Oxford University Press

Thompson KG & Schall JD (1999) The detection of visual signals by macaque frontl eye fields during masking. nature neuroscience 2(3):283-8

Todd J J & Marois R (2004) Capacity limit of visual short-term memory in human posterior parietal cortex. Nature 428, 751-754

Van der Stigchel S, Heslenfeld DJ and Theeuwes J (2006) An ERP study of preparatory and inhibitory mechanisms in a cued saccade task Brain Research 1105: 32-45

Varela FJ (1996) Neurophenomenology: a methodological remedy to the hard problem. Journal of Consciousness Studies 3:330-359

Vogel E K & Machizawa M G (2004) Neural Activity predicts individual differences in visual working memory capacity Nature 458:748-751

Wang X-J (2001) Synaptic reverbrations underlying mnemonic persistent memory. Trends in Neuroscience 24:455-463

Wolpert DM & Ghahramani Z (2000) Computational principles of movement neuroscience. Nature neuroscience 3:1212-7

Wurtz RH & Sommer MA (2004) Identifying corollary discharge for movement in the primate brain. Progress in Brain Research, vol 144, chapter 3, pp47-60

Xu Y & Chun M M (2006) Dissociable neural mechanisms supporting visual short-term memory. Nature 440:91-5

Zahavi D (1999) Self-Awareness and Alterity. Evanston Ill: Northwestern University Press

Zahavi D (2005a) Being Someone. PSYCHE 11(5):1-20

Zahavi D (2005b) Phenomenological Approaches to Self-Consciousness. Stanford Encyclopedia of Philosophy.

Zahavi D (2006) Subjectivity and Selfhood: Investigating the First Person Perspective. Cambridge MA: MIT Press

Glossary

A-Consciousness: Action consciousness of Block's 2-factor theory of consciousness

ADAN: Anterior-directing attention negativity

CERN: Centre European de Recherche Nucleaire in Geneva, Switzerland

CODAM: Corollary Discharge of Attention model of consciousness creation in the brain

EDAN: Early-directing attention negativity

EEG: Electroencephalography

ERP: Event-related potential

FEF: Frontal eye fields, in prefrontal cortex

FC5/6: Fronto-central placements for EEG electrodes

fMRI: Functional magnetic resonance imaging

GW: Global workspace (basic to Baar's model of consciousness in the brain)

HOT: Higher order thought model of consciousness

IPS: Superior intraparietal cortex

LDAP: Late-directing attention positivity

LHC: Large hadron collider at CERN

MEG: Magnetoencephalography

N1-P1: Complex of ERPs at 80-150 msecs post-stimulus

N(T): Target negativity

N2pc: Posterior contra-lateral brain activity during the N2 phase of post-stimulus brain activity

N2pca: Early parietal-based component of the N2pc, from 180-220 ms post-stimulus

N2pcb: Later temporal-lobe-based component of the N2pc from 240-280 ms post-stimulus

P-Consciousness: Phenomenal consciousness in Block's 2 factor model of consciousness

PC4: Post-central placement of electrodes in ERP measurements

P(D): Distracter positivity

PET: Position emission tomography

PFC: Pre-frontal cortex

PO7/8: Parieto/occipital placing of electrodes in ERP measurements

Ptc: Positivity sited contralateral activity in temporal lobe

RLIP: Right-lateralised inhibitory potential

SPCN: Sustained posterior contralateral negativity

SU(3)xSU(2)xU(1): The standard model of the elementary particles, with the separate symmetry groups SU(3) for the strong/nuclear interactions, SU(2) for the weak/radioactive interactions and U(1) for the electromagnetic interactions

TMS: Trans-cranial magnetic stimulation

V1-n: Visual processing modules in the occipital cortex of the brain

VSTM: Visual short term memory

WM: Working memory (as short term memory)

W & Z Bosons: Particles, discovered at CERN, whose exchange creates the weak force of radioactvity