Distribution Agreement

Theo Young 16 April 2013

Relational Perspective Neutral Monism: an Alternative Approach to Establishing a Science of Mind/Consciousness

by

Theo Young

Dr. Richard Patterson
Adviser

Department of Biology

Dr. Richard Patterson

Adviser

Dr. Alexander Escobar

Committee Member

Dr. Arri Eisen

Committee Member

2013

Relational Perspective Neutral Monism: an Alternative Approach to Establishing a Science of Mind/Consciousness

By

Theo Young

Dr. Richard Patterson

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Department of Biology

2013

Abstract

Relational Perspective Neutral Monism: an Alternative Approach to Establishing a Science of Mind/Consciousness

By Theo Young

This paper explores the problem of consciousness by attempting to fit the mind within a naturalistic framework of the universe. Contemporary neuroscience has made enormous progress in understanding brain function and mechanisms from a behavioral standpoint, but one problem continues to elude us: how the brain, a three pound organ, can be conscious – not just a processor of sensory input, but, as Thomas Nagel wrote, “that there is something it is like to be” an organism with a brain (Nagel 1974). Until now we have mostly conducted science based on an approach that considers ourselves as external observers examining the universe, but it is time to put ourselves, the conscious observers, back into the universe. No paradigm of reality can be considered complete unless it incorporates consciousness. In order to achieve a naturalistic understanding of consciousness it will be necessary to modify the way in which we view reality, not as a universe of physical matter, from which somehow the mental derives, but as a neutral reality, of which the ‘physical’ and ‘mental’ are just different representations from different perspectives.

Relational Perspective Neutral Monism: an Alternative Approach to Establishing a Science of Mind/Consciousness

By

Theo Young

Dr. Richard Patterson

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Sciences with Honors

Department of Biology

2013

Acknowledgements

I would like to express enormous gratitude for my adviser Dr. Richard Patterson, for his helpful advice and wisdom, and his patience in guiding me through writing this thesis.

Also, a tremendous thank you to Dr. Arri Eisen, for without his approval for my – rather unorthodox – topic for biology, I would not have been able to write this paper.

Many thanks to Dr. Alexander Escobar, who encouraged me to be bolder in my thesis and expand my topic beyond mainstream neuroscience.

Finally, a special thanks to Dr. Eric Weeks for helping me understand the double-slit experiment that demonstrated wave-particle duality.

Table of Contents

Abstract	1
Introduction and Overview	1
Chapter 1	3
Chapter 2	10
Chapter 3	14
Chapter 4	18
Chapter 5	23
Chapter 6	26
Chapter 7	28
Chapter 8	35
Chapter 9	44
Chapter 10	50
Chapter 11	58
Chapter 12	65
Chapter 13	81
Chapter 14	82
Chapter 15	85
Figures	86
References	102

List of Figures

Figure 1.1. Binocular rivalry, used to test neural correlates of visual awareness. 86

Figure 1.2. Logothetis binocular rivalry experiment. 87

Figure 1.3. Meditation and gamma synchrony. 88

Figure 1.4. Location of the posteromedial cortices (PMCs) in the brain. 89

Figure 3.1. Relational Perspective Neutral Monism (RPNM). 90

Figure 5.1. Tononi's multidimensional qualia shape, or "quale. 91

Figure 7.1. Three stages to developing a science of mind. 92

Figure 7.2. Actual translative relationship between the 'mental', 'physical', and neutral. 93

Figure 8.1. Cross-communications between the PMCs and other areas of the brain. 94

Figure 10.1. Reconstruction of visual experiences from brain activity. 95

Figure 12.1. The double slit experiment that demonstrated wave-particle duality. 96

Figure 12.2. Bell's Theorem. 97

Figure 12.3. Wavefunction collapse provides possible translative laws. 98

Figure 12.4. The difference between first and third person perspective representations. 99

Figure 12.5. The von Neumann-London-Bauer-Wigner and Penrose versions of the wavefunction collapse are just different ways of describing the same phenomenon. 100

Figure 12.6. Hameroff's microtubule model for quantum computation. 101

“The promise of future science is to furnish a unifying goal to mankind rather than merely the means to an easy life, to provide some of what the human soul needs in addition to bread alone.”

– Eugene Wigner

RELATIONAL PERSPECTIVE NEUTRAL MONISM: AN ALTERNATIVE APPROACH TO ESTABLISHING A SCIENCE OF MIND/CONSCIOUSNESS

Theo Young

Abstract

This paper explores the problem of consciousness by attempting to fit the mind within a naturalistic framework of the universe. Contemporary neuroscience has made enormous progress in understanding brain function and mechanisms from a behavioral standpoint, but one problem continues to elude us: how the brain, a three pound organ, can be conscious – not just a processor of sensory input, but, as Thomas Nagel wrote, "that there is something it is like to be" an organism with a brain (Nagel 1974). Until now we have mostly conducted science based on an approach that considers ourselves as external observers examining the universe, but it is time to put ourselves, the conscious observers, back into the universe. No paradigm of reality can be considered complete unless it incorporates consciousness. In order to achieve a naturalistic understanding of consciousness it will be necessary to modify the way in which we view reality, not as a universe of physical matter, from which somehow the mental derives, but as a neutral reality, of which the 'physical' and 'mental' are just different representations from different perspectives.

Introduction and Overview

Consciousness is difficult to tackle scientifically due to a curious problem. The scientific method is tailored to and has thus far only investigated phenomena from third-person ontology but consciousness comes to be known from an exclusively first-person experience. Once the subjective is objectified, the subjective aspect is lost, which is what was trying to be explained in the first place. There are two distinct contemporary neurobiological approaches to solving the problem of consciousness: trying to break qualitative experience down to its components, and developing models to describe consciousness from a unified point of view. Both however, are flawed because they are looking for the physical basis of the mental. This is based on relational perspective confusion, a misunderstanding of what the 'mental' is. Neurobiological approaches, as they are currently formulated, will only get better correlations and more accurate predictive models, but will not explain consciousness, because the mental is not a 'thing' that is created or caused by physical brain processes, nor is it identical to specific physical processes, but rather it is a perspective, or point of view. This paper proposes an alternative paradigm of mental-physical relation – relational perspective neutral monism (RPNM), which provides a more coherent viewpoint as to how exactly the 'mental' fits within reality. According to RPNM, the mental is not caused by physical events, but rather 'mental' and 'physical' are two different perspectives, first and third person, respectively, of the same 'neutral' reality – neutralities, or neutral events – which we do not directly access through our perception. Rather, it is accessed via sensory inputs and internally represented. The accuracy of this representation depends upon the evolutionary benefit it provides the organism. Spatial representations are thus likely to be highly accurate, but the neutrality that is consciousness in the first person is not practical to be represented as such in the third person, so is simply represented as objects external to the self, 'brains'. Within RPNM, it will be necessary to expand scientific inquiry beyond just the third person perspective to include first person as well; rather than establishing how the mental arises from the physical, the problem becomes translating between these two perspectives, first and third person. To do so, we will need to understand this neutral reality (i.e. we will need explanatory laws translating between neural event – neutrality – mental event). With RPNM, is a scientific explanation of consciousness still viable? One possibility is that we cannot understand this neutral reality because we cannot directly access it because of the way our perspective works, in which case an explanatory science of consciousness will not be achievable. However, a predictive model based on establishing correlations between the first person 'mental' and third person 'physical' is certainly attainable using the current methods in neuroscience, and can be considered a science of mind. This outcome, though effective, is rather unsatisfying. A second possibility is that the neutral reality can be conceptualized and understood, in which case an explanatory model of consciousness will be possible. The current neurobiological approach does not seem to be taking us in this direction. However, ideas in more fundamental domains of science, including those that currently remain on the fringe, are making an attempt to go beyond correlations to understand this deeper reality. In order to understand the neutral reality we must incorporate evidence from these areas as well. Ultimately, it is necessary to keep an open mind to the possibility that our conception of reality may need to be expanded in order to include the mental.

1. Contemporary Neuroscientific Approaches

The mainstream neuroscientific approach within the science of mind has two goals: first, to break down qualitative experience to its components by establishing specific neural correlates, and then to develop models from that evidence to describe consciousness from a unified point of view. It would be impossible to cover all of the models emerging from this approach here, but I will highlight those which are relevant to the focus of this paper.

Neural correlates of consciousness (NCC) were proposed by Francis Crick as a means of finding specific differences in the neural activity of a subject when he is conscious versus when he is not (Crick 1994). Most experiments conducted on neural correlates involve isolating neural correlates of specific aspects of consciousness. The most well known of these were the binocular rivalry tests conducted by Nikos Logothetis. Binocular rivalry describes a phenomenon that occurs when a subject is presented with a different visual stimulus in each eye (Fig. 1.1). For example, the left eye can be presented with a face image and the right eye with a flower image. Instead of consciously perceiving both images simultaneously, the subject can only be conscious of one or the other image at any given time. After a few seconds the perception stabilizes into an alternation between the two images, being conscious of one image for around five seconds, then the other for five seconds, and so on. Logothetis et al. examined the specific neural correlates of conscious perception of faces using binocular rivalry by simultaneously stimulating monkeys with a face and non-face image (Leopold and Logothetis 1999). The monkeys were trained to indicate when they perceived a face via pushing a lever. The monkeys' brain activity was monitored by microelectrodes measuring action potentials of individual neurons located within various visual cortical areas: the striate cortex (V1), extrastriate areas (V2 and V4), middle temporal area (MT), medial superior temporal sulcus (MST), inferotemporal cortex (ITC), and the upper and lower bank of the superior temporal sulcus (STS, known to become active when presented with visual face stimuli). It was found that certain individual neurons responded only when the monkey reported *conscious perception* of the face, despite the binocular visual stimulus remaining constant throughout. Overall, the microelectrode measurements of the visual cortex neurons indicated that 90% of ITC and STS cells, 40% of V5/V4 cells, and less than 20% of V1/V2 cells responded when perception change was reported, but that the other percentage remained active as long as the face visual stimulus was present (Fig. 1.2). This implies that the primary visual cortex's role is mainly in detection of visual stimuli, but that the higher visual cortices are involved with conscious perception. V1 activity remains largely the same since the visual stimulus does not change, but the conscious experience alternates between face and non-face perception as the higher visual cortical activity changes correspondingly.

Rather than narrowing focus on correlating specific brain areas with specific aspects of awareness, other studies have been conducted to establish more brain-wide neural correlates of general consciousness. Perhaps the most well recognized brain-wide correlate of consciousness is synchronous gamma wave oscillation, which was established using electroencephalography (EEG). When a single visual stimulus causes receptive fields not directly connected/communicating with each other to become activated, the neurons within those fields synchronize and oscillate at 25-100Hz, the gamma range, with typical frequencies at around 40Hz (Gold 1999). For this reason Crick and Koch postulated that gamma synchrony may play a role in the 'binding problem' of visual awareness, how different visual stimuli/inputs can come together in the brain to create a unified visual experience. Additional experiments with anesthesia showed that gamma synchrony is selectively terminated when a subject undergoes general anesthesia, indicating a broader connection with general consciousness and not just visual awareness (Schwender et al. 1994). Recent EEG studies with Tibetan Buddhist monks revealed very intriguing links between meditation and gamma synchrony (Lutz et al. 2004). In a normal meditative state, the monks' level of gamma wave oscillations was similar to that of a control group of untrained meditators. However when both groups were instructed to meditate upon thoughts of "unconditional loving-kindness and compassion", the monks exhibited a very strong brain-wide neural synchrony of gamma oscillations from 25-42Hz, with the highest amplitude ever measured in mentally healthy human brains (Fig. 1.3). Additionally, the monks' higher frequency gamma waves from 80-120Hz also exhibited a significant increase in amplitude compared to that of the control group. The monks also displayed a much more defined brain-wide, long range global gamma synchrony in their bilateral frontal and parietal/temporal regions than the untrained meditators. This evidence indicates that higher levels of consciousness (during meditation, as experienced meditators regularly report being 'more aware' during deep meditative states) are associated with higher levels of gamma synchrony.

A few neuroscience models have attempted to create a unified framework of the mind to describe what consciousness actually is. According to Antonio Damasio, the construction of a complex 'sense of self' and its relation to the external world is essential for consciousness (Damasio 2010). As neural networks became more complex, animals sequentially evolved three states of the constructed self: proto, core, and autobiographical. The protoself entails neural representation of the body in stable state, accompanied by feelings of body in said state. The core self is a brief united sequence of modified protoself images (occasionally accompanied by feelings) that arise from a disturbance of the protoself by an 'object(s)'. Generation of the autobiographical self entails what human beings regard as 'consciousness.' The autobiographical self is created when a large sequence of core self images are coupled into a coherent pattern. Thus in order for the brain to generate the autobiographical self, it must first possess both a proto and a core self.

Damasio believes the posteromedial cortices (PMCs, known commonly by its component parts: the posterior cingulated cortex, retrosplenial cortex, and precunneus) are essential in mediating coordination and cross communication between different cortical areas (Fig. 1.4), which is vital in constructing an autobiographical self (more details on Damasio's model later, when I discuss self-to-external environment representation and its implications in how we perceive reality). Studies in anesthesia, sleep, behavioral neuroscience, and coma/vegetative state research have revealed compelling evidence linking PMCs activity with consciousness, particularly 'self-consciousness.' From functional magnetic resonance imaging (fMRI) studies, it was determined the anesthetic propofol acts primarily on three sites: the PMCs, thalamus, and brain-stem tegmentum – all major players in Damasio's neural model for construction of the self. Of particular interest, during use of propofol and other anesthetics, the decrease of blood flow to the PMCs is highly correlated with the decrease in level of consciousness in the patient (Alkire and Miller 2005). During non-rapid eye movement sleep (N-REM, 'dreamless'), when consciousness levels are at their lowest, brain-stem, thalamus and PMCs activity are also decreased significantly compared to the awake state. During rapid eye movement sleep (REM, 'active dreaming') sleep, consciousness levels are slightly elevated, and so is PMCs activity (though still less than in waking state). In a series of positron emission tomography (PET) and fMRI studies, it was demonstrated that PMCs activity is highly elevated during performance of tasks concerning self-reference, but diminished during tasks extensively engaging external stimuli (Damasio 2010). Finally, a few patients in a vegetative state do gradually recover, and as metabolic activity increases in their brain-stem, thalamus, and most predominantly, PMCs, so do they slowly regain consciousness as well.

Giulio Tononi's Integrated Information Theory is one of the more distinctive approaches to modeling consciousness. Tononi postulates that at the fundamental level consciousness is integrated information, the degree of which he denotes with the Greek letter 'phi' (Tononi 2008). But what exactly is 'integrated information'? Tononi uses two thought experiments to illustrate the concept.

For information, imagine a simple photodiode placed in front of a screen that can detect if the screen is 'light' or 'dark'. Now imagine a human in front of the same screen, who is instructed to say 'light' when light, and 'dark' when dark. Both the 'conscious' human and the 'unconscious' photodiode are performing essentially the same function, so what is the difference between them? The difference is how much information is generated by the system performing the action. Information, according to Tonini, is the reduction of uncertainty: the more alternatives that are ruled out, the greater the reduction of uncertainty and the greater the information. When the photodiode is detecting the light it only generates two possible alternatives, light and dark. However, a human being detecting the light involves a much higher number of possible alternatives. Suppose for example, instead of just a blank screen of 'light' or 'dark' the screen played a movie. The photodiode would not be able to distinguish what was going on and would continue to report light or dark based on its criteria for determining between the two alternatives. A human being however, would be able to understand the complex images on the screen. Even when a human is differentiating between light and dark, he is generating a wide range of alternatives to 'light' and 'dark'. The photodiode is not, and cannot even discriminate between the different shades of light; past a certain threshold all light is the same. Additionally, it does not even know that it is detecting light and dark, and not hot and cold, for instance. A photodiode's version of light is only one of two possible things it knows, but our version of light is one of an innumerable amount of different possible experiences – this gives it a much different meaning for us than for a photodiode. Our ability to distinguish pure light from all these alternatives makes us more conscious. The more alternatives there are, the more specific the definition of 'light', and thus the more conscious one is of 'light'.

So why is integration needed? Imagine a camera, whose detectors distinguish around 2^1000000 alternatives, corresponding to 1 million photodiodes or bits of information. Why then is the camera not conscious like the human brain, which has a similar number of bits of information? Or consider the cerebellum, which contains many more neurons, and thus more information, than the cerebral cortex. Why then do cerebral cortical lesions result in loss of consciousness whereas cerebellar lesions do not? The answer lies within the degree of integration within the system. Even though the camera system can distinguish between 2^1000000 alternatives, it has no way to integrate the 1 million photodiodes; they cannot communicate with each other. Likewise, the cerebellar neurons are nowhere near as interconnected as the cerebral cortical neurons. In reality the camera system is just a collection of 1 million isolated photodiodes; each photodiode distinguishes between only two alternatives independently of every other photodiode. If the camera chip were divided into each one of its 1 million photodiodes, its capability would not decrease. This cannot happen with a brain. In an integrated system, the phenomenological experience is that of a whole, and cannot be broken down into separate parts – a red square cannot be split into the experience of red, and of a square, separately. Splitting the corpus callosum may separate the left and right visual fields, but it creates two separate (each unified in and of themselves) consciousnesses. The elements in the brain work together as a unified, integrated system, which ceases to work when separated, unlike the camera. So to be conscious, a system has to distinguish between many alternatives – information, and must be an integrated system.

IIT carries a curious implication. There is a slight problem with the statement that consciousness is limited to mammalian (or even animal) brains. If we attribute subjective experience to one thing and not the other, why so? What does that former thing have that the latter does not, so that we can attribute consciousness to it? One might say complexity, or level of integration/cross-communication between the integrated elements within a system. But integration and cross-communication only mean a more complex system, in no way do they imply the sudden magical generation of an inner subjective experience. Most importantly, we cannot describe a complex integrated system in relation to a simple non-integrated system in a way which allows us to independently infer that the former is conscious, but not the latter. With

IIT we are left with the inevitable conclusion that every system is conscious to a certain degree, by our new definition of 'consciousness'. There is no magical threshold of consciousness; it comes in degrees. Our brain is a system which is conscious to a high degree, but each individual neuron in the brain is also a less conscious system, each organelle, organic molecule, atom is also a conscious system. Likewise, on the macro scale, each planet, solar system, star cluster, galaxy, and so on are also conscious systems. Of course, they are all systems with a relatively low phi in comparison with the human brain (or any animal brain), so there is no need to assume that the Milky Way Galaxy, for example, has the same kind of experiences we do.

2. Why Have All Neurobiological Attempts to Solve the Problem of Consciousness Fallen Short Thus Far?

With these recent advances in neuroscience, there is no reason to doubt that at some point we will be able to understand how consciousness is realized in the physical brain, or is there? Something unique about consciousness makes it unlike any other subject of scientific inquiry. Approaching the problem of consciousness through neuroscience, from a third person empirical perspective has thus far always led to a dead end, but is there any reason for this other than 'the problem is very difficult but the complex mechanisms will be worked out eventually'?

Contemporary views of the mind are divided between two general factions: those who think physical neuroscientific models cannot explain consciousness, and those who do. Both sides are faced with their own version of the 'problem of consciousness'. Defenders of the 'hard problem' generally espouse one or more variants of the same age-old claim, "I can know every possible thing about the physical brain but I will never find consciousness." Within contemporary neurobiology, which approaches the problem from a physicalist view of reality, a slightly different problem exists, "What is the physical basis for the mental?" Both of these positions stem from the same historical misinterpretation of what the mental (and physical) actually are. They are the result of a perspective confusion, combined with an inadvertent tendency to envisage the 'mental' and 'physical' as 'things' – a vestigial habit from the days of dualism that still lingers within our linguistic descriptions and conceptualizations.

The 'hard problem' question can be approached by asking, "What is consciousness?" It is not just the process of aware observation of one's surroundings in relation to oneself, but also one's internal thoughts, feelings, subjective sensations – 'qualia', if you will. Consciousness, in other words, *is* the first person perspective. The first-person perspective is not just critical to, but *defines* consciousness. Consciousness cannot be empirically observed from the third person; how exactly is the 'first-person perspective' supposed to be represented as such through third person observation? It simply cannot be done – first person perspective cannot be observed except *from* the first person perspective.

How could one identify another's 'first person perspective' from the third person point of view? What would it look like? Could you see it, touch it, hear it, smell it, taste it? The idea is logically incoherent. This scenario is reminiscent of the famous exchange that Ludwig Wittgenstein had with his student Elizabeth Anscombe, in which he asks, "Why do people say it was natural to think that the sun went around the earth rather than the earth turned on its axis?", to which she replies, "I suppose because it looked as if the sun went around the earth." Wittgenstein then remarks, "Well what would it have looked like if it had *looked* as if the earth turned on its axis?" (Anscombe 2001). Wittgenstein is obviously implying that it would have looked the same as if the sun revolved around the earth, but his analogy reflects a poor conception of relational perspective. If it had looked as if the earth turned on its axis, it would look like *the entire globe, being viewed by me from outer space, rotating on its axis*. The point of reference would be entirely different from that when it looked like the sun revolved around the earth. This problem of relational perspective is not only why we used to think the sun revolved around the earth, it is also why we have been confused about the nature of consciousness for so long. Applying Wittgenstein's allegory to consciousness:

Q: Why do many people think there is a 'mind-body problem'?

A: Because we can dig around the brain all we want and we won't find consciousness. Based on our scientific paradigm there is no evidence to suggest it exists.

Q: Well what would it look like if consciousness *does* exist?

Obviously to a third person observer it would look *exactly* like if consciousness did not exist, because consciousness by its very definition *is* 'the first person perspective'. But as with the earth/sun analogy, there is more to the answer. If consciousness does exist, it would look like *me, as a first person observer, experiencing and relating other entities/my surroundings to myself*. Once again, the point of reference would be entirely different from that when it looked like consciousness did not exist. Just as we would have to be in outer space to see the earth rotating on its axis, consciousness can only be known of introspectively; it *is* the first person perspective. This is why there is so much confusion arising with respect to these issues; we never – or at least rarely – consider picturing the situation from a different perspective. For millennia we thought the sun revolved around the earth, and some still think there is a 'mind-body problem' because we cannot 'find consciousness' in the brain, all because of a relational perspective confusion. Consciousness is not a 'thing' with a physical basis, but rather just the first person perspective. 'Consciousness' events are represented just like all other events, as other physical entities relative to the 'organism self' – in this case, as 'physical brain' events. It could *never* be any other way. So of course we will never 'find consciousness'; how do you expect to find ‘first person perspective’ from third person observation? We are attempting to objectify subjectivity itself. The moment we define consciousness objectively, we lose the subjective aspect of it, which is what we were attempting to pin down in the first place. The statement ‘I cannot find consciousness no matter how much I study the brain!’ is about as earth-shattering as the statement ‘I cannot see the earth rotating upon its axis no matter how much I study the sunrise/sunset!’ The self-evident truth of the aforementioned statements do not imply any sort of revolutionary metaphysics, but is simply a result of the way perspective works, how an organism relates other entities/its surroundings to itself.

The ‘neural/physical basis for consciousness’ problem, like the ‘hard problem’, derives from the same unintentional dualist influence; all descriptions of ‘consciousness’ as a ‘thing’ that is ‘created/produced/generated’ ‘by’ or ‘in’ the brain are tremendously misleading and generate much confusion. ‘How does the mental arise from the physical?’ is not a scientific problem; it is a logical misunderstanding rooted in our perspective dependent representations of reality. Though few today remain substance dualists, the Cartesian dualist *way of thinking* still dominates the way we conceptualize reality. Although using dualist terminology to split reality into the ‘mental’ and ‘physical’ proved helpful with traditional third person scientific empiricism, it is impeding progress in the science of mind. Neither ‘physical’ nor ‘mental’ really exist; they are imaginary boundaries we have set based on how we relate ourselves to the external world. Simply put, ‘physical brain events’ is just how an organism represents ‘mental conscious events’ in the third person, as ‘another entity/surroundings’ relative to itself. ‘Mental conscious events’ are the first person perspective of the same events. In actuality they are identical, and the confusion arises because they seem to be different events due to a different frame of reference. Physicalism is essentially just substance dualism sans the ‘mental’. If we continue to picture reality as ‘physical’, with everything deriving from it, the quandary of how mental events can be caused by (or identical with, if one is applying identity theory – more on this later) physical brain events will always exist. ‘Mental’ and ‘physical’ are descriptive terms, not metaphysical entities. Both original dualist thinking and the elements of it that still creep into our folk psychology are the result of a linguistic confusion, which itself stems from a reference point confusion. ‘Mental’ and ‘physical’ are linguistic terms we use to describe the two different relational perspectives we have of the same thing. We have built a misconception of the ‘mental’ and ‘physical’ as distinct ‘entities’ because of our continued usage of these terms as nouns; we have a predisposition to conceptualize ‘perspectives’ as ‘things’. Perspectives are not ‘things’; they are points of reference.

3. Relational Perspective Neutral Monism, an Alternative Paradigm to Physicalism

So what then *does* exist? In my model of relational perspective neutral monism – RPNM (Fig. 3.1), reality is composed of a neutral reality/substance, a ‘neutrality’ that is neither mental nor physical. In this case the label ‘neutral’ is not used to assign any sort of metaphysical properties to reality, it is simply a way to shake off the confusing and misleading terms ‘mental’ and ‘physical’. Since we cannot directly access this neutral reality, we should presently refrain from making assumptions about its properties. In this case assigning the term ‘neutral’ to reality absolves it from any preconceived notions about what it must be like. According to RPNM, there is one reality, just different perspectives of observing/looking at it. The ‘physical’, as we currently define it, is not reality as it truly is, but rather a perspective dependent representation of reality. We do not have direct access to the neutral reality, because rather it is accessed via inputs from the five senses and reconstructed within the brain. Thus, depending on how helpful the representations are in aiding an organism, an organism would have evolved to represent reality accurately in some respects but not so much in others. The first and third person perspectives exist as a result of the way we have evolved to represent reality, as a self relating to the external world. In the case of the 'brain', what is represented as a physical brain in the third person is consciousness in the first person perspective. In order to translate between the two perspectives it will be necessary to understand the nature and properties of the neutral reality from which both perspectives are derived. Right now, for simplicity, I will use physicalist terminology (e.g. 'brain' or 'neurons' instead of 'neutral elements', even though a 'brain' is just a third person representation of a neutrality) to describe RPNM, since describing in terms of the neutrality would inadvertently assign properties to it, the nature of which we do not know.

There is often a lack of agreement as to what actually constitutes the 'mental'. While this is true, as far as I can tell, no one has come up with an adequate definition of the 'physical' either. RPNM gives a clear and unambiguous definition of both: 'mental' is the first person perspective representation of a neutral event and 'physical' is the third person perspective representation of that same neutral event.

As the terms 'first person' and 'third person' can be used interchangeably depending on what frame of reference one has, it is important here to clarify what these terms mean, and how they are used within the context of RPNM. Depending on what frame of mind one has at any given moment, he can conclude either one of two things. First, that the third person representations of reality (physical) are exclusively part of the first person perspective. This is known as *reflection*, that though corresponding to an external reality, everything we experience is actually generated internally by our brains. Dreams demonstrate that we can even internally construct an 'external world' in the absence of actual external stimuli. We smell, hear, touch, see, and taste in our dreams, perhaps not as vividly as in waking life, but they are all the same types of experiences. If one dreamt about conducting an experiment, the process of observation would not change, but the 'third person empirical observation' in the dream would be completely a result of first person representation. Such it really is with waking life as well. The difference is that in dreams my internal representations do not correspond to an external *represented* reality that other individuals can represent as well. It is purely a product of activity within a single brain. One could say it is representing in the absence of a represented external reality. There is nothing being *represented*, only one integrated system (brain) *representing*; thus it is entirely first person. When taken to an extreme, this idea leads to solipsism.

Conversely, we can consider our third person representations to be an accurate depiction of reality as it truly is. This is known as *pre-reflection*, and is a less introspective, more evolutionary primitive frame of mind (animals navigate around their world assuming that what they experience is true reality, not just an internal neural representation of reality). This idea may lead to extreme versions of materialism, such as behaviorism, which deny the existence of mental states, since we cannot empirically observe consciousness in the third person. There is no need to accept either of these extremes. We have good reason to believe that third person representations, although comprising part of first person experience, are not entirely concoctions of the aforementioned, but do correspond to actual external realities. There are evolutionary reasons as to why this should be true, as I will discuss later, but one of the most immediately obvious reasons is because we cannot mentally control what goes on in our third person representations of the external world, unlike during lucid dreaming. However, we do not need to assume that these third person representations are always an accurate depiction of the external reality; that version of reality – as physical matter – at least leaves out the first person perspective/consciousness. As for the existence of the internal reality, the first person perspective, it is self-evident. Descartes was correct about at least one thing.

According to RPNM, consciousness is an integrated neural system constructing a self-map in relation to its external environment. Similar neural systems, as part of the external environment for a particular individual 'self', would be represented as 'physical brains'. Essentially, integrated neural systems represent each other, so 'third' and 'first' person, like 'subjective' and 'objective', are nothing but more linguistically misleading terms. However they are necessary in the context of explaining the RPNM position.

For convenience, a good definition of/distinction between first and third person perspective is this: 'third person' describes the representation of 'neural system A' which every other system except 'neural system A' itself will observe, and 'first person' describes the representation of 'neural system A' which is only observed by 'neural system A', and by no other system. For example, in the case of my neural system, I will observe it as a first person experience, an ego/self in relation to the external world; the first person perspective. Every neural system other than my own will observe it as a physical brain; the third person perspective. Naturally, this definition is a bit sloppy, since strictly speaking there is no guarantee *any* representation of an external event is the same for all systems. In other words, it is true that no one can know of my experience of pain, which we generally label as a first person experience, but it is also true that no one can know what I see when I observe the results of an experiment, which we generally attribute to the third person objective. This solipsistic uncertainty actually does not matter in practice, because although theoretically all our representations of the external reality could differ individually, in third person empirical study (science) we are still able to reach a consensus as to how the external events operate. In other words, in third person observation it *appears* that we all represent the same external events in the same way, and that is adequate. This does not hold true for what we generally consider as first person experiences: no one can come to any consensus as to how each individual's experience of pain operates.

It comes down to this: the neutral 'system' *doing the representing* is the first person perspective, and the system(s) *being represented* is/are the third person perspective. When the system is doing the representing, it is experienced in one way (the first person perspective); when the same system is being represented, it is experienced in a different way (the third person perspective). Obviously, a neutral system cannot be doing representing and being represented by itself at the same time, which is why we are incapable of experiencing our own 'brains' as we experience other 'brains', as a three pound conglomeration of neurons and glial cells.

Thus for practical purposes, and within RPNM, when something is labeled as a 'third person representation', it refers to when we consider the representation to be part of the external reality, and when something is labeled as a 'first person perspective representation', it refers to when we consider the representation to be part of our inner experience. For example, when speaking of an apple on a table, we can consider it to be either part of the external reality (in pre-reflective mode), or a product of my inner conscious experience (in reflective mode). In reality of course, it is neither: there is no 'apple', as we represent it, in the external reality, but the manifested image 'apple' in my conscious experience is not a total product of my imagination. It does correspond to a real thing; it is a representation of a neutrality in the external reality, by another neutrality (of which the third person representation would be 'brain').

4. Why was Neutral Monism Unnecessary in Other Domains of Science?

Here I will explain why traditional physicalism has worked for Western science so far, and why it cannot work for the science of mind. Traditional physicalists tend to oversimplify the issue by stating that our third person representation (of the external surroundings) is an accurate depiction of reality based on the successes of Western science thus far, concluding that the physicalist approach to studying the mind will prove fruitful. However it is not difficult to see the obvious problem that arises with respect to first person perspective. The 'neutral substance' which is represented as 'physical matter' in the third person, can be assumed to be and described as such in other scientific disciplines, because in those domains we are investigating external events/phenomena and their relations with each other. Physicalism thus works for this 'neutral substance' without any problems in other scientific domains, which exclusively employ third person empiricism. In other words, we investigate everything in these other domains from the same perspective, the third person, so the problem of perspective dependent representations never arises. However when it becomes necessary to incorporate consciousness, this physicalist model hits a brick wall. Consciousness *is* the first person perspective, which cannot be observed except *from* the first person perspective. To sidestep this dichotomy, physicalists will attempt to redefine consciousness as a 'thing' with a physical basis (not as a perspective dependent representation: the first person point of view), and search for its physical cause. This of course is bound to fail. The reason neutral monism is not considered in any other domain of science is because there is no need for it – there is no change in perspective. There was no problem in any other scientific domain because our perspective never changed, we always conducted our empirical observations from the third person perspective, so everything we observed we managed to classify together based on our third person representations of all of them, constructing a coherent third person representation of reality. The scientific method has thus far relied on the third person perspective, observations from that point of view. However, studying consciousness shifts perspective from third to first person. This is a clear dichotomy that is unique to consciousness alone. With respect to consciousness, this is the first time we have to consider the first person perspective, the first time our investigation involves a change in perspective. If we don't involve the changing of perspective, the incorporation of first person perspective, we will be stuck with attempting to explain how the mental 'stuff' arises from physical matter. Physicalism ultimately ignores the first person perspective and constructs the entire scientific paradigm of reality around our third person representations of the world. Although metaphysical dualism has fallen out of favor, it is imperative to shake off the dualistic way of conceptualizing and linguistically describing the 'mental' and 'physical' as well. In order to progress in science of mind, we must acknowledge that neither 'physical matter' nor 'mental consciousness', as we define them, are the metaphysical foundation of reality, or even real at all; both are just concepts and human constructs, ways in which we represent the neutral substance that *is* the foundation of reality. Concepts evolve around our frame of reference, and language evolves from concepts. 'Physical' and 'mental', are ultimately just linguistic terms.

Some physicalist neuroscience models incorporate identity theory: rather than attempting to explain *how* the 'mental' arises from the 'physical', they opt instead to make identity claims between specific neural events and specific mental events. Once a specific neural correlate of consciousness is established, an identity claim is drawn. For example, neural mechanism 'A' would be identical to mental event 'A' in the same way that water is really identical to H_2O. These models are slightly superior to those attempting to discover a causal mechanism for the generation of consciousness from physical brain events, as they do not carry the erroneous assumption that consciousness is a physical 'thing' created by the brain. However, as Tom Nagel noted, identity theories still neglect to explain why brain events and mental events, if identical, can be so phenomenologically different (Nagel 2012). Why for example, would a specific pattern of neural activity be the same thing as an experience of blue color? The reason has to remain within the confines of physicalism, and yet no ‘physical’ property of neural events is able to draw a link between it and the phenomenological experience. It is possible to imagine the experience of blue color existing independently from the neural activity, but not possible to imagine water existing independently of H_2O. Furthermore, a comprehensive explanation of water allows one to understand *how and why* H_2O is identical to water, but a comprehensive model of neural correlates of consciousness does not allow one to understand *how and why* specific mental events are identical with specific neural events. A complete descriptive framework of H_2O based on its physical properties alone would allow one to independently deduce its identity with water, but a complete descriptive framework of neural correlates of consciousness based on their physical properties alone would not allow one to independently deduce their identity with subjective experiences, unless an unjustified identity claim is explicitly made. Many current ‘unified models’ of consciousness employ variations of the identity claim. For example, in Integrated Information Theory Tononi argues that integrated information *is* consciousness (Tononi 2008), but neglects to explain how this identity claim can be drawn. Other ‘unified models’ seem to avoid (or at least sidestep) specifically addressing consciousness altogether. For example, Damasio’s autobiographical self model does provide a hypothetical descriptive model for the consciousness we know a priori to exist, but ultimately fails to actually explain how the construction of an autobiographical self would entail a conscious organism rather than just a non-conscious automaton with a highly accurate self navigational system.

The assumption that our third person perspective representation of external reality is always accurate is a vestige from the days when it was taken for granted that the Divine had endowed us with the capacity to experience the world as it truly was. It was this assumption that allowed us to proceed and accomplish so much with the scientific method, grounded in third person empiricism. Our current concept of the physical leaves out the mental, and vice versa, so how can either of them be reality as it is? Both are incomplete pictures, perspective dependent representations of reality. We must shake off the intuition that reality is all physical, that our third person perspective of reality must be reality as it truly is. The idea that the physical must be true reality is additionally absurd because my third person representation of the 'physical world', as I indicated earlier, actually constitutes part of my first person representation. So making an identity claim and saying that the qualitative experiences in the brain are just 'neural mechanisms' is rather ridiculous because the third person representation of 'neural mechanisms' itself constitutes a qualitative aspect of a first person perspective. We do not have Godlike direct access to reality; all our concepts of reality are based on frame of reference, and representation. Accepting RPNM would be completely compatible with all existing scientific knowledge, but additionally, unlike before, there would now be no problem as to how the mental arises from/is identical with the physical. In RPNM there exists no problem with epiphenomenalism, consciousness plays an active role in the universe, and it operates under what we refer to as 'the laws of physics'. The 'neural mechanism that causes my hand to move to pick up the glass of water' is the third person representation of the neutral event which is represented as 'my conscious decision to move my hand to pick up the glass of water' in the first person perspective.

With physicalist neuroscientific approaches to consciousness based on identity claims one can achieve better correlations and more accurate predictive models but can never reach an explanation of why specific neural events are identical with specific mental events. But the scientific method is not necessarily physicalist, and can be reformulated to operate within a different paradigm of reality, in this case, RPNM. When these formerly physicalist neuroscientific models are applied instead within the context of RPNM, the necessitation that our perception of the external world always accurately depicts reality is done away with. Consequently the problems of how consciousness is produced by the physical, or how mental events can be identical with brain events (because they are *not* identical, but rather both derived from the same neutral reality), disappear, but our current scientific method is preserved, and one can continue with the actual science of building a framework for this discipline.

5. Applying Neuroscientific Models within the RPNM Paradigm

Replacing traditional physicalism/materialism with RPNM will both rid us the problem of consciousness *and* be entirely compatible with the scientific method and all our current scientific descriptions of the world. Here are two examples of how existing neuroscientific models can integrate into RPNM. Take for example, olfactory perception. Without going through the specific details, the basic pathway of olfaction involves odor molecule ligands that bind to receptors on the dendrites of olfactory receptor neurons, which then send action potentials via their axons (which comprise the olfactory nerve) to the mitral cells in the glomeruli of the olfactory bulb, which eventually transmit the signals to different areas of the brain, notably the piriform cortex (Leffingwell 1999). Where then does the first person experience of a specific smell fit into this? A traditional physicalist framework does not explicitly address it, or may simply make an identity claim between the very phenomenologically different 'subjective smell' and the neural signaling pathway. RPNM provides a much clearer perspective: the aforementioned neural signaling pathway is the third person representation of the neutral events which are represented as the qualitative perception of specific smells in the first person perspective.

How would RPNM look within a unified model such as Tononi's? Integrated Information Theory at first glance, unlike other neuroscientific models of consciousness, seems to be immune from the physicalist trap and appears not to even need to be redefined within RPNM. Could integrated information after all be the neutral element RPNM seeks? It seems plausible; Tononi claims that integrated information *is* consciousness, and information is not a physical entity. However, this statement is a form of, or at least analogous to, identity theory. When Tononi equates integrated information with consciousness, he is speaking within the third person frame of reference, assuming a third person representation depicts reality as it truly is. It is very important to clarify here that *the third person perspective representation does not necessarily have to be 'physical'; anything described in objective terms falls into the same trap: it lacks the subjective aspect*. Thus his description of 'integrated information' is still an incomplete representation of reality, and he attempts to mediate that by drawing identity statements between it and qualitative experiences. If Tononi truly believes that integrated information is an accurate depiction of the reality as truly is, he should expect much more predictive power from this model. For example, Tononi uses "*quale*" or multidimensional qualia shapes as a configurative depiction of qualitative experiences (Fig. 5.1). These shapes are modeled based on neural cross-communications. Each point in the multidimensional shape symbolizes a specific qualitative aspect. So qualia shapes similar in structure are expected to produce similar unified qualitative experiences (Tononi 2008). This provides a great deal of predictive power; we can see how it would be possible to predict the qualitative nature of novel experiences based on the similarity of their corresponding 'qualia shapes' to known 'qualia shapes'. Even so, this could only be accomplished by means of assigning identity statements between specific 'qualia shape' characteristics and specific subjective qualitative aspects, based on observing specific neural events correlating with specific qualitative experiences. We would still have no way of bridging the gap between third and first person, of translating the 'qualia shape' into the qualitative experience. For example, consider a deaf man, who has never heard any sound before. In his case he cannot predict the qualitative experience of novel 'qualia shapes' by associating these shapes with 'qualia shapes' of already known experiences, because he has no known experiences (with respect to hearing). So if he were to look at the 'qualia shape' of middle C played on a clarinet, would he be able to know what middle C on a clarinet sounded like without ever having experienced it beforehand? Obviously not, which is why, as useful as the IIT model is, it still does not bridge the gap between third and first person. Unless an identity claim is made, a comprehensive explanation of IIT will not allow one to deduce the nature of subjective experiences, or even that consciousness exists at all. RPNM rephrases the identity statement to reflect the actual situation: 'integrated information is a third person representation of the neutral reality that is experienced as consciousness in the first person perspective'. So the 'qualia shape' of middle C is a third person representation of a neutral event, which is represented as the experience of hearing middle C in the first person perspective.

To respond to criticism about subjective states not being deductible from a description of IIT, Tononi states that "being is not describing". He uses this analogy: *understanding* how nuclear fission works does not *make* nuclear fission occur; likewise *understanding* how consciousness works does not *make* consciousness occur. This is true, but consciousness presents us with a unique problem. There is something very curious about consciousness as opposed to everything else. With consciousness you *have* to *be* in order to know what it is like. In every other science, you do not have to *be the process* to know what it is like; you can know just from the description of the process. Once again it is the perspective difference problem. If integrated information *was* consciousness, we could describe IIT to someone and he would be able to conceive of the first person just from understanding IIT. As it stands, this is not possible. In the case of nuclear fission (Tononi's analogy), I can describe it to someone, and although his imagination of nuclear fission is not equivalent to nuclear fission actually occurring, it would still be an accurate representation of what happens when nuclear fission *does* occur. He would be able to imagine what nuclear fission was *like* just based on its description. Can the same be said for consciousness? If I describe IIT to someone, and he imagines IIT, will that be an accurate representation of what happens when consciousness occurs? No, because what he imagines is nothing like what is going on from the first person perspective. He would have to *be* in the first person even just to know *what it was like*; he would not be able to imagine it just based on the description of integrated information alone. This problem never arises in any other science. So there are still two perspectives in this case, which never occurs in any other science. The nuclear fission analogy is thus slightly misleading.

So Tononi should not make the mistake of *equating* integrated information with consciousness. We must realize that we only describe it as integrated information because we, as of now, can only explain things in terms of our third person representation of reality. The concept of integrated information still leaves the first person perspective out.

6. Why is RPNM Necessary for Progress in Science of Mind?

At this point, the reader is probably wondering what the practical difference between physicalist identity theory and RPNM really is. If one cannot make unjustified identity claims between a specific neural event and a specific mental event, how does the application of a neutral reality help the situation? It may be true that RPNM is compatible with existing neuroscientific models and approaches to consciousness, but what does it contribute, how can it further progress in science of mind, why is it necessary? Simply rephrasing “brain event A is identical with mental event A” into “brain event A is the third person perspective of the neutral event that is mental event A in the first person perspective” does not contribute anything apart from the introduction of an unnecessary third ‘neutral event’. It certainly leaves us no closer to a solution to the problem of consciousness; the gap between mental and physical still remains, and we are still left with nothing but correlations. However, RPNM *does* provide a possible means of going beyond correlations, which existing neuroscientific models based on identity theory do not. The furthest identity theory can reach is building a correlative model by drawing identity statements between specific neural events and specific mental events, but with no real explanatory framework of how the two are actually identical. In other words, it still has not bridged the explanatory gap between physical and mental. Attempting to establish direct identity laws between the physical and mental will be unsuccessful because these laws will necessarily be confined within physicalism, and thus derived from the physical ‘third person’ properties of neural events, which cannot allow us to deduce the first person perspective. However, RPNM is not confined to physicalism; *through* the neutral reality we can establish translative bridge laws between the mental and physical. If somehow we were able to understand the neutral reality in its entirety, we would be able to deduce both the first and third person perspectives that are derived from it. Once this is accomplished, the explanatory gap between the mental and physical can be bridged and it will be possible to translate between the two perspectives. How exactly is this to be done?

If we assume that brain events are related to mental events via the neutral reality in a way such that bridge laws may eventually be drawn between them, we will need to construct a framework that translates a physical neural event A into a mental event A. In this scenario, we can make an analogy between the proposed ideal 'science of mind' and our current scientific understanding within the discipline of genetics. In order to construct a complete science of mind in Scenario A, one with soundly accurate predictive power, we must develop a framework which allows us to *understand why* 'neural event A' translates to 'mental event B'. Establishing a correlation between specific neural events and specific mental events, no matter how precise, will simply not suffice.

7. Three Stages to Establishing a Complete Science of Mind

Here I will propose a three stage program to help us potentially reach an explanatory science of mind, using an analogy drawn with the science of genetics (Fig. 7.1). What is the difference between an explanatory scientific framework and one grounded in establishing correlations? Suppose that in a hypothetical development of the science of genetics, early geneticists had established that DNA was indubitably associated with the presence of proteins. After decades of meticulous study, the geneticists announce that every specific protein has been correlated with a specific nucleotide sequence in human genome. Would we have achieved a true science of genetics? Absolutely not; geneticists would have no understanding of *why* a nucleotide sequence results in the presence of a specific protein. They would be completely ignorant of transcription, RNA processing/export, translational polypeptide synthesis, and protein folding, all the elaborate steps between a nucleotide sequence and the final protein product which provide an explanation as to *how* the former gives rise to the latter. If the science of genetics constituted merely establishing correlations, geneticists may be able to identify a specific protein with every *known* nucleotide sequence (and vice versa), but what if a completely novel sequence was introduced? The geneticists would have no means of predicting the protein associated with that sequence. Likewise, if neuroscientists established a correlation between every known neural event and every known mental event, they would still be unable to predict any mental event caused by a completely novel neural event. Nevertheless, this still provides a good launching pad for establishing a scientific framework of consciousness. Consider it Stage One.

Modern geneticists *can* predict the specific protein* associated with any novel sequence, due to our deciphering of the genetic code. However, one must not be fooled into thinking that cracking the code constitutes a complete science of genetics. In this case we have simply established an algorithm for translating a specific DNA triplet codon into a specific RNA triplet codon into a specific amino acid. It is essentially still a framework based on correlations, albeit a very accurate one that has managed to break down nucleotide and protein sequences into discrete components, triplet codons and amino acids. We still do not necessarily understand the biochemical processes that are involved in transcription, RNA modification/export, and protein synthesis. Likewise, neuroscientists may be able to construct an algorithm for 'translating' every neural event 'codon' into its corresponding mental event component/qualitative aspect, in which case any mental event could be predicted given any novel neural event. This is Stage Two. This can be established with a framework based on very accurate correlations that break subjective experiences and brain activity into very discrete and specific qualitative aspects and types of neural signaling, respectively. Using the earlier example of a middle C played on a clarinet: a specific type of neural signal can be linked to the qualitative aspect of the timbre of a clarinet, and another neural signal to the qualitative aspect of the pitch middle C. When both these neural signals are simultaneously active, we can determine that the qualitative experience of a middle C played on a clarinet occurs. In this case, analogous to cracking the genetic code, novel qualitative experiences *can* be predicted by assessing novel neural signaling activities. However, it would still be an *artificial translative* algorithm, and we would still be ignorant to the nature of the *actual translative* underlying bridge laws between a neural event and the resulting mental event in the real world. Because of this, the deaf man in the above example still cannot predict the qualitative sound experiences of novel neural events by associating these events with neural events of already known experiences, because he has no known experiences (with respect to hearing).

Only when the *actual translative* underlying bridge laws have been understood, will we know *how and why* a neural event actually translates into a mental event, and only then will individuals having total lack of known experiences be able to nevertheless predict the qualitative nature of an experience corresponding to a novel neural event. In the science of genetics this has been essentially accomplished; we are able to trace the pathway (transcription → RNA processing/modification → RNA export → translation/polypeptide synthesis → protein folding/assembly of subunits) a nucleotide sequence takes to become a protein. To successfully construct a science of mind, we must understand the underlying laws that govern the translation between a neural event/third person perspective and a mental event/first person perspective. This is Stage Three. This can only be attained by acquiring the translative bridge laws between the neutral reality and our third and first person representations of it, and from there deriving the bridge laws between the two perspectives by seeing how they relate to each other via the same neutral reality. Therefore, unlike the previous two stages, which do not involve the neutral reality, Stage Three will require a comprehensive understanding of the neutral reality itself.

Following the aforementioned modus operandi to creating a science of mind, Stage One is to establish correlations between the first and third person perspectives. The traditional empirical approach that has worked for everything else in science so far – assuming that the third person perspective provides a complete and thoroughly accurate representation of reality – will not work here because it leaves out the first person perspective which is imperative to science of mind. To successfully construct a science of mind, it will be necessary to expand scientific inquiry beyond the third person perspective to include the first person as well. However, one cannot simply represent the 'first person perspective' as 'first person perspective' from third person observation. In order to observe the 'first person perspective' as 'first person perspective' we must *be* in the first person, as in 'me as a self experiencing my surroundings'. So since we cannot access another's first person perspective, only our own, a successful science of mind hinges upon gathering information by means of introspective techniques. This approach has thrived in Eastern philosophy in their unique schools of idealist thought. In the 1920s a similar approach known as phenomenology, introspectively examining consciousness from the first person perspective, was initiated in Western philosophy by the writings of Edmund Husserl (Phenomenology 2013). It is important however not to rely solely on introspective techniques for information gathering; an ontology based *purely* on first person introspection will leave out the physical, just as pure physicalism leaves out the first person perspective. We would be wrong if we said 'consciousness does not exist' based on our third person observations of the brain, but we would also be wrong if we said 'the mind is all that exists' based on our first person introspective observations. In both of these cases we're missing the whole picture. Scientific empiricism has relied solely on third person observation, but Eastern philosophy has a long history of involving first person introspection. In order to effectively develop a science of the mind, the first stage would be to incorporate elements of Eastern introspection and/or Western phenomenology within the scientific discipline. This is being done already.

Neurophenomenology is a new but developing branch within neuropsychology, and seeks to combine information gathered via introspection with third person empirical observations in neuroscience. The late neurobiologist and philosopher Francisco Varela, who encouraged the incorporation of Buddhist introspection into the science of mind, was perhaps the most well-known pioneer in neurophenomenology (Varela and Shear 1999). Varela was a strong proponent of using highly trained Buddhist meditators as subjects for neuroscientific study, as well as training neuroscientists themselves in Buddhist meditative techniques. His rationale was that when it came to studying consciousness, it was much more reliable to study Buddhist meditation practitioners, who were highly trained in the art of reproducing, not objective phenomena as in Western science, but specific subjective experiences, than ordinary laymen who were not trained in these techniques. Much like it would be more practical and reliable in any other scientific domain to use someone highly trained in a particular field to replicate specific experimental results, rather than an untrained layman. Once we have constructed a basic framework correlating general aspects of subjective awareness with respective neural events, we will be on our way to establishing a science of mind. However, it is important *not* to stop here, like the identity theorists, and create a neural correlate framework based purely on identity claims.

Stage Two is to establish an artificial translative algorithm for translating a highly specific type of neural activity into a specific qualitative aspect. A successful Stage Two model will be able to predict the mental events associated with novel brain events, based on a highly accurate correlation between specific qualitative aspects and specific neural activity. However, it will still not be able to achieve a true understanding of *how and why* a certain brain event translates into a certain mental event. Stage Two has not been accomplished yet, but there is no reason to see why a successful framework of this type will not be forthcoming. A number of models of consciousness are working towards this goal, such as the ‘qualia shapes’ in IIT, where specific types of structures within various qualia shapes correspond to specific types of qualitative aspects in subjective experience.

Stage Three is to understand the actual translative underlying bridge laws between a neural event and the resulting mental event in the real world. For the sake of concision I sometimes refer to these laws as translative laws between physical and mental events, but to be more precise, as I noted earlier, since both 'mental' and 'physical' are different representations of the same neutral reality, the laws must actually be based on translative laws between the neutral reality and both the first person 'mental' and third person 'physical' perspectives which mutually derive from it (direct bridge laws between mental and physical imply identity theory). Actually the translative laws relating 'mental', 'physical', and neutral are even slightly more complicated (Fig. 7.2). As I noted earlier, my 'third person' representation of a particular neutrality actually constitutes part of my first person perspective. This means a translation between a specific first and third person perspective needs to involve two separate neutralities. For example, say we wanted to translate between the first and third person perspective of 'neutrality A'. The first person representation of A is 'an experience of the color red' – call it 'mental event A', and the third person representation of A is 'some neural activity between the visual cortex and the ventral stream' – call it 'brain event A'. We could use first person translative bridge laws to directly translate between neutrality A and the first person 'mental event A'. However, we cannot use third person translative bridge laws to directly translate between neutrality A and the third person 'brain event A'. This is because the 'third person' representation of neutrality A, as 'brain event A', actually constitutes a *first person representation of a separate neutrality B* – call it 'mental event B' – within the observer's own brain. [†] So to translate neutrality A into its third person representation we first must use third person translative bridge laws to translate neutrality A into the corresponding neutrality B that is activated when the observer looks at neutrality A, *then* use first person translative bridge laws to translate neutrality B into *its* first person representation 'mental event B'. In this case, the first person representation of neutrality B, 'mental event B', would be identical to the 'third person representation' of neutrality A, 'brain event A'. ‡

Once we have established these translative bridge laws, we will be able to see the relation between the first person aspect that derives from a neutral property, and the third person aspect that derives from the same neutral property. It is then that we will be able to draw the connection between correlating first and third person, 'mental' and 'physical' events, and translate between the two perspectives. To find the translative laws between the neutral reality and the first and third person representations of it, we will obviously need to understand the neutral reality. This is the most precarious stage. Can it even be accomplished? At this point there are two possibilities. One possibility is that because of how perspective has evolved within an organism relating itself to the external environment, the neutral reality can never be accessed and therefore fundamental bridge laws translating between the physical/third person and mental/first person will never be established. Another possibility is that although we are unable to directly access the neutral reality, we can still develop a scientific model that accurately conceptualizes it, thus achieving the elusive bridge laws. I argue that an effective and practical science of mind is still reachable regardless of which possibility is true.

*Not really, as our knowledge of the biochemical mechanisms involved in peptide folding are still limited, but this is not relevant to the point I am making.

†This is another reason why identity theory is incoherent. Drawing an identity claim between a specific 'brain event A' and a specific 'mental event A' does not work because both are in fact 'mental events', corresponding to

different neutralities! ‘Brain event A’ is identical to a ‘mental event B’ corresponding to a neutrality B, and ‘mental event A’ corresponds to a neutrality A.

‡Some eliminative materialists such as Daniel Dennett have argued that qualitative aspects themselves are the result of a misguided folk psychology, and do not exist in actuality, because our various conceptions of qualia, i.e. necessary characteristics we associate with specific qualitative aspects, do not accurately correspond to the way physical brains organize sensory information (Dennett 1988). But based on my model above, our perception of others’ ‘physical brain function’ (third person representation) is also a qualitative aspect (first person representation) within our own brains! So if we are to accept Dennett’s argument that qualitative aspects do not actually exist, than it follows we will not be able to trust our observations – or rather, qualitative experiences – of a subject’s brain function.

8. A Possible Hindrance from Reaching Stage Three

First, why is it that we cannot directly access the neutral reality? The reason lies within how perspective works, how animals have evolved to represent the external world. I have thrown the word ‘representation’ around very liberally, but how exactly do mammalian brains ‘represent’, and what are the implications this has for how we see the world, and consequently, how science is conducted and its possible limitations? The external world must be ‘translated’ in some way by the integrated neural system, but does that ‘translation’ necessarily reflect reality? Many neurobiologists, such as Tononi, make the argument from natural selection to support the position that consciousness accurately depicts the external reality as it truly is (Tononi 2008). However, I only agree to a certain extent. Natural selection does strongly imply that extreme solipsism is wrong; our representations correspond in some way to an external reality. It does not imply however that our representations will *always* accurately depict external reality as it truly is.

Many experiments have been conducted to determine how brains represent the external world and construct a self-external relationship map. Packard and Teather for example experimented with spatial representations in rats (Packard and Teather 1998). Rats were placed in the bottom arm of a cross shaped maze, and trained to learn that food was always located at the end of the left arm of the maze, relative to the bottom arm. When the rats were placed in the top arm however, instead of turning left, they turned right to coincide with the location of the food, now at the end of the right arm of the maze relative to their new location. Further tests revealed that if the hippocampus is lesioned, the rats will lose their spatial orientation and exhibit a trained conditioned response, consistently turning left relative to themselves regardless of where in the maze they were placed. These experiments imply that rather than having behavioral responses for local directional navigation, such as turning right or left, programmed within the brain, mammals by way of the hippocampus are able to construct highly accurate spatial representations based on mapping the self in relation to the complete external environment.

How are external inputs coded by neurons? One possibility is that a specific qualitative aspect is coded by a single neuron, known as local coding (Churchland 2002). Another possibility is scalar coding, in which the firing rate of the neuron codes for variations in qualitative aspects. Most likely however, qualitative aspects are coded for by varied firing patterns in a collection of neurons, with overlap in the features coded for by each individual neuron, known as vector coding. Local coding is reminiscent of a run of the mill, oversimplified identity theory, in which every qualitative aspect has its corresponding specific single neuron; for example, the color blue has a specific ‘color blue neuron’ located on the receptor sheet of the retina. Vector coding is much more complex; specific qualitative aspect are not coded for by single neurons, but rather patterns of neural activity within populations of neural networks. Individual neurons in the population will contribute to the coding by varying its firing rate. A vector is a set of numbers symbolizing the firing rate/activity levels of individual neurons within the population. Similar firing patterns/vectors by a population will code for similar qualitative aspects; for example, vector {17,8,9} might code for aquamarine and vector {17,9,10} might code for cyan. Vector coding allows a smaller number of individual neurons the ability to code for a much larger repertoire of qualitative aspects. Consider a population of 5 neurons with four levels of activity, from 0-3. In local coding, only 20 features (5x4) can be coded, but in vector coding, 625 features (5^4) can be coded.

Once the external inputs are processed, how do the mammalian, specifically primate, brains use that information to construct a self to relate to the external environment? According to Antonio Damasio's autobiographical self generation model, to go about constructing a self, the brain requires tools capable of retrieving memories from dispositional cortices (in the cerebral cortex), treating them as biographical 'objects' to make them conscious in 'core pulses', and coherently interacting them with the protoself in the brainstem (Damasio 2010). What structures in the brain have the capacity to facilitate this broad scale coordination? Certainly the job cannot be delegated to a single structure, but rather involves controlled collaboration between several structures. Damasio proposes several possible candidates, each of which may contribute to assembling the autobiographical self. The thalamic nuclei are considered due to their location halfway between the cerebral cortex and the brain stem, allowing for cross communication between the two areas. The most serious contenders according to Damasio are the series of interconnected cortical areas he terms "convergence-divergence regions (CDRegions)". These include the temporparietal junction, the lateral and medial temporal cortices, the lateral parietal cortices, the lateral and medial frontal cortices, and the posteromedial cortices (PMCs). Other possible coordinators are the claustrum (a thin layer of neurons sandwiched between the insular cortex), the putamen, and the thalamus, considered for their connections to various sensory regions and their distinct forward (but weak backward) projections to the PMCs.

Of the aforementioned candidates, Damasio believes the PMCs contribute most significantly in constructing the autobiographical self, and consequently, consciousness. Neuroanatomical studies have revealed that the PMCs receive converging inputs from a myriad of brain regions, including the parietal and temporal association cortices, entorhinal cortices, frontal cortices, anterior cingulated cortex, claustrum, basal forebrain, amygdala, thalamus (intralaminar and dorsal), premotor region, ventromedial prefrontal cortex, and frontal eye fields. The PMCs also provide output signals back towards these regions (with the exception of the ventromedial prefrontal cortex, claustrum, and intralaminar thalamic nuclei). Damasio hypothesizes that the PMCs act as a 'higher-order CDRegion' which integrates a wide range of information distributed within other CDRegions (Fig. 8.1). The PMCs, like the thalamic and brain stem nuclei, are bilaterally symmetrical and located near the brain midline, consistent with Damasio's postulated 'shared traits' of consciousness-related structures.

Once the self is generated, how do brains accurately relate it to representations of the external environment? Successful eye-motor coordination and movement with relation to the external environment (as well as the organism's own body), requires a region of the brain that is able to translate sensory information (primarily visual and auditory) into accurate relational representations relative to the 'self' located within the head behind the eyes. The brain needs to distinguish between the organism's own body and external objects and accurately map the location of these elements relative to each other. The posterior parietal cortex plays this role, translating sensory inputs into a self-external relationship map that facilitates motor coordination to allow the organism to interact with its environment. Pouget and Sejnowski's experiments have shown that area 7a and 7b of the parietal cortex play a major role in eye-hand coordination and spatial representation (Pouget and Sejnowski 1997). When area 7 is lesioned in monkeys they exhibit difficulty in relating objects in space, faulty eye movement and impaired navigational skills. Area 7 is multimodal, meaning it is comprised both of cells that respond exclusively to visual, auditory, somatosensory, chemical vestibular, or proprioceptive signals, and multimodal cells that respond to several types of signals.

All this evidence points towards the generation of a coherent 'self-to-external environment' relationship map by the brain. Animals have evolved to represent reality in a way which makes them function effectively. Our evolved methods of third person representation translate external neutral reality's characteristics into qualitative characteristics that are most beneficial with regards to the organism's survival. Anything beyond this is unnecessary. (Why are our qualitative experiences of electromagnetic radiation as heat and visible light so different? They are actually the same phenomenon.) . Sometimes it just so happens that reality *is* represented as accurately as possible, and sometimes it is not, but it all depends on what the functional purpose of representations requires in any given situation. In this case, to navigate within the environment without running into trees or falling off cliffs, our third person representations for most of the external world, such as location of obstacles, threats relative to the self and depth/spatial representation, are likely to be quite accurate (although these representations can be manipulated occasionally via illusory techniques, such as ventriloquism) – these representations must correspond to the external world in such a way that navigating with our self-external relationship map we do not inadvertently put ourselves in harm's way.

But how accurately do brains represent specific qualitative aspects of the various external elements? Qualia exist because animals require a means of distinguishing different external elements from each other, e.g. a buffalo is not the same thing as a lion. Additionally, neurobiological evidence suggests that as we become more familiar with a particular subject, our qualitative representations of elements within that subject will become increasingly complex and refined (Churchland 2002). For example, most humans have no problem recognizing distinct facial features of fellow human beings but find great difficulty in distinguishing individual members of a population of chimpanzees. However, researchers who have spent many years studying chimpanzees will easily be able to recognize and distinguish unique facial features of each individual chimpanzee. The visual input of the chimpanzee face for all human beings is identical, so what has changed? The higher visual centers involved in integrating the visual information and neurally recreating it as a represented image have remapped themselves in the case of the experienced chimpanzee researcher.

What is the significance of this with respect to representing other 'minds'? As we have seen, qualia exist as a means of distinguishing external elements, so qualitative aspects do not necessarily indicate anything about the nature of the neutral property that is actually being represented. Representing other 'minds' as 'first person perspectives' would make little sense in terms of constructing a coherent self-external relationship and it makes sense that animals did not evolve with this ability. If we were to accurately represent the first person aspect of all reality, the self-external relationship map would fall to pieces; there would exist multiple selves simultaneously that would have no way of relating to each other. In order to have a coherent, biologically functioning organism, it must construct a relationship map consisting of *one* self, relating to everything else – the non-self 'external reality'. Thus all other 'selves' are represented not as first person perspectives, but as non-self 'things' external to '*my* self'. The qualitative aspects of other minds are only represented insofar as to distinguish them from other external elements. Thus animals evolved to represent them as various neural events in physical brains, like any other 'non-self' aspect of the external world. This is one obvious case where evolution has thwarted our capacity to observe and measure a certain aspect of reality. Because of how we evolved, our third person representation of the external world by necessity leaves out the first person perspective. Everything else we describe in science pertains to the external we observe relative to the self. However, when with the same traditional scientific approach we use on everything else, we attempt to describe *the very thing doing the observing* (not the 'physical brain', as that is how we represent someone else's observing system, as an external element relative to the self, but the actual personal conscious self, the first person perspective that is observing everything external relative to it), we arrive at an impermeable barrier.

Conversely, how does the brain represent itself? My own integrated system, brain, cannot represent itself in the third person, as a physical brain. The brain can only represent itself as the first person perspective, the ego/self experiencing everything in relation to itself: the physical body and its sensations, as well as everything external to the physical body. Certain things we feel like are 'outside' relative to the self: pain, visual stimuli, smell, sounds, and certain things we feel like are 'inside' relative to the self: feelings of joy, sorrow, confusion. These latter things we cannot pinpoint to a specific location relative to the self. This qualitative aspect of these representations says nothing about the actual relation of these events relative to the representing brain. As phantom limb experiments demonstrate, a subject can feel a pain outside the self in a specific part of the body that is no longer in existence (Ramachandran and Hirstein 1998). The representation of pain is a complete product of the representing brain. Usually it does correspond to an actual event external to the brain, but – as the phantom limb phenomenon illustrates – not always. So our own brain, or integrated system, cannot represent itself in the 'physical brain' manner not because it is metaphysically impossible, but because we did not evolve to do so. If our own integrated systems did represent themselves as ‘physical brains’, then the integrated system that constructs the self would be representing itself as an object external and relative to the self, which would cause complete confusion, as well as disruption of homeostasis and organismal function.

In order to have a coherent self-external environment relation, we have evolved to split our representations of reality into two different types: mental (first person) and physical (third person). Neutralities either comprise the self (a single integrated system, such as the brain) or are represented as something relative to the self (everything else not a part of the aforementioned integrated system). How we represent neutralities is a result of what is most evolutionarily beneficial for us as organisms. It is a result of our attempt to construct a coherent self-external environment relation. The concepts ‘first person’ and ‘third person’, and consequently ‘mental’ and ‘physical’, arise out of the evolution of a self-external relation process. So is reality a self (first person ‘mental’) or something external, relative to the self (third person ‘physical)? Neither – all our concepts of reality are the evolutionary result of an organism building a coherent, functional, effective self-external environment map/relationship. In order to effectively do this, the integrated system has to represent itself and the same type of integrated systems external/relative to itself differently/in different ways, as first person ‘mental consciousness’ and third person ‘physical brains’, respectively. This means an accurate representation of reality is not evolutionarily plausible, because if reality were represented accurately, as neutralities from which both first person ‘mental’ and third person ‘physical’ can be derived, a functional self-external relationship map would be impossible.

What is the implication of all this? It certainly indicates we do not have direct access to the neutral reality, but we do have access to our representations, which, in order for the organism to function effectively, are not necessarily accurate depictions of reality. In order to achieve a successful self-external reality relation, there *must* be different representations for the same type of neutrality (e.g. the representations 'physical brain' and 'conscious self/ego' correspond to the same neutrality), and we can only observe something in either the first or third person perspective, never both simultaneously. For the first time in our history, we are faced with a very obvious setback of our evolutionary derived means of perception. It is biologically impossible for us to directly observe or measure the neutral reality, as we do phenomena in other domains of science. I can only access a subject's 'neutral events' by observing them in the third person, as physical brain events. In order to observe them as a first person perspective, I would somehow have to *become* the subject and experience from *his* first person point of view. One possible consequence of this is that the gap between the first and third person points of view can never be bridged; we are biologically incapable of accessing the neutral reality our two perspectives derive from, and thus are unable to understand how one translates into the other.

So the colorblind neurobiologist Mary in Frank Jackson's knowledge argument will never be able to experience the color red, deaf neurobiologist Frank likewise will never be able to know what Bach sounds like, and Laplace's Demon[§] will never be able to predict consciousness, even if they understand everything about the brain, because none of them have access to the first person perspective of these neutral events. They are like the astronomer who does not have access to a satellite – he will never be able to *see* the earth rotating on its axis even if he knows everything about the earth's rotation by studying the sunrise/sunset (he could be able to *imagine* what it looks like but this is irrelevant to this analogy, since the neurobiologists cannot even imagine what the first person perspective of these neutral events is like, since they have nothing to base their imagination off of – the astronomer at least can base his imagination on his knowledge that the earth is round, and its relation to the other planets and the sun).

§Jackson's 'Colorblind Mary' knowledge argument and Laplace's 'Demon' thought experiment both argue that one could know everything about the brain but still not be able to predict the qualitative experiences of consciousness (Jackson 1982). These conclusions however are self-evident within RPNM, because they essentially state, 'No matter how much I study something from one perspective I will never be able to experience it as I would experience it from *another* perspective, when I am looking at it from *this* perspective.'

9. Why An Effective And Practical Science of Mind Can Still Be Constructed Even if Stage Three is Not Accomplished

If this is the case, how can we ever hope to construct a science of mind? Essentially, a 'traditional' science of the mind, one that explains how and why physical brain events translate to mental conscious events cannot be established, since the neutral reality from which both perspective dependent representations are derived cannot be directly accessed/observed. Effectively, we cannot go beyond Stage Two of establishing an artificial translative algorithm.

However, this can still constitute a practical and useful science of mind. Although we cannot bridge the gap between the first and third person perspective, with Stage Two we can still construct, based on neural correlates of consciousness, a highly accurate predictive framework that correlates specific brain events (third person) with specific mental events (first person), and even predict novel mental events given novel brain events. Of course, we would have to acknowledge that any 'brain event' derives from the same neutral event as its corresponding 'mental event', despite not being able to scientifically bridge the gap. This raises an interesting question: if the gap cannot be bridged, how then will we ever be able to definitively ascertain the truth of RPNM? Unfortunately, if it is indeed true that the gap between first and third person perspective cannot be bridged, RPNM *cannot* be definitively proven; all we will ever get is correlates. However, that model will still be successful considering its aims; it will be a comprehensive scientific model in the same sense as in any other scientific domain, because most scientific theories work through just this type of approximative application. How so?

Our understanding/conception of the world is built around our self-external relationship map and is divided into two perspectives: first person and third person. Science is a tool we use to describe our observations within this self-external relationship conception of reality, but it does not necessarily need to reflect a comprehensive understanding of the fundamental reality. Physicist David Bohm used a simple analogy to illustrate our inability to grasp the true nature of reality (Bohm 1990): suppose we were studying an ellipse. Upon first inspection it seems the ellipse is the appearance and the circle is the essence. But then we realize that the circle was an appearance as well; it is mostly empty space, and is made of atoms. Then we find that the atoms are made of protons, neutrons, electrons, such that the atoms were just an appearance. Then we find that subatomic particles are also an appearance, because they are actually made of elementary particles, and so on (or according to Tononi, reality is in fact none of the above, but essentially integrated information!). It becomes an infinite regression that keeps on receding but we never get to the true essence of reality, because everything we observe is based on perception, how we represent reality by contextualizing/compartmentalizing it within our self-external relationship map. We evolved to represent reality within a self-external relationship map (everything not a component of the integrated system ‘self’ is represented as the ‘external’/‘physical’, the first person perspective is excluded from all other integrated systems not within the integrated system ‘self’) because that way was the most organismally beneficial. Thus we are biologically incapable of studying consciousness from the third person point of view, so extending science of mind beyond an accurate neural correlates framework may not be possible. Although most obvious in science of mind because of the change in perspectives, this extends to other scientific disciplines as well, as I noted above. Each time we think we come closer to finding the 'essence' of reality, it is simply another representation made within our self-external relationship maps, which does not depict reality as it truly is. As long as we are stuck within our self-external relationship maps, we will never have access to the true 'essence' of reality.

If this is so, how can a scientific theory ever be considered to be correct? Bohm remarked, "Fundamentally, science is involved in a perceptual enterprise…if an appearance is 'correct' it is in some way related to the reality but it is evidently not the reality." What we deem to be 'correct' in science is simply something that allows us to construct a coherent, useful model of that particular phenomenon we are describing. It is an extension of our general representations of what is 'external to the self' in a way which allows us to function effectively as organisms. None of these representations, be they ways in which we see the world from day to day, or complex scientific models, are perfectly accurate depictions of true reality; they are simply useful ways in which we represent reality that gives us predictability, so that we can function efficiently. According to Bohm's definition of 'correct', IIT (or another predictive model) could be the 'correct' theory of 'consciousness', depending how accurate its predictions are. We make no necessary distinction between a 'correct' scientific theory and the true nature of reality in other scientific domains because as I stated above, there is no perspective change, but in the science of mind it is obviously necessary because two different perspectives, first and third person, are both involved here.

Scientific 'laws' are descriptions we assign to predictable phenomena that we observe, to help us make future predictions (e.g. there is no 'essence' to a magnetic field, it does not exist as a thing it is just a paradigm/model we assign to this behavior that we observe to help us make future predictions). They may apply only within a specific scenario under certain circumstances, but sometimes they can cross over and 'bridge' different 'levels' of reality/domains.

According to RPNM everything reduces to a fundamental neutral reality in which a unifying body of laws will not just be domain-specific approximations, but metaphysically true statements about the actual nature of reality. This reality is not accessible, which may mean that we will not be able to see the end product of that reduction, as the nature of our perception entails an infinite regression as we reduce reality to a smaller and smaller scale; we can never grasp true reality. Stage Three may never be achieved. Physics is the not the end of our reduction; it is not the foundation of reality. When we approach the Planck scale, at the quantum level, even the elegant Newtonian laws begin to break down. We have no reason to think that quantum laws are fundamental, and they will likely break down as well once we go to an even smaller scale. Recently the Higgs particle was confirmed, but immediately the question was raised, "What is that made of?" The endless regression continues to open up novel domains with new 'laws' governing ever smaller levels of reality. Thus all our scientific laws in the various domains, economics, psychology, biology, chemistry, Newtonian mechanics, quantum physics, integrated information theory, are just approximations (albeit very useful and accurate approximations within their specific domains) we selectively use depending on which scenario, scientific domain/level of reality we want to make predictions in. However, none of them get us to the true reality.

In all practicality however, it is unnecessary to reduce all domains to their constituents (e.g. reduce psychology to neuroscience or biology to physics) when we have very good approximative ‘laws’ which allow us to make astoundingly accurate predictions within each specific domain/’level’ of reality. In this case, different neurobiological models of the mind, as long as they demonstrate high levels of accurate predictive power, can be considered as good candidates to employ in a science of mind. It does not matter that they do not provide an explanatory bridge between the first and third person – they need not be fundamental truths, but like the laws in every other scientific domain, only highly accurate descriptions/approximations.

This does not mean that science is useless, or even that any arbitrary model can be liberally applied within the science of mind. Jessica Wahlman argues that because all scientific domains are approximations or ‘narratives’ as she calls them, and because purely objective third person empiricism cannot investigate subjectivity, we should turn to “our social skills and sympathies as much as (if not more than) a pre-determined set of physical criteria. In order to come up with useful explanatory tales, our, literary imaginations and moral sympathies will be an indispensable asset in constructing sound theories” (Wahlman, 2013). However, this seems little better than guesswork; although it is true that our scientific ‘laws’ are domain specific ‘narratives’ we use to help us make predictions within particular scientific disciplines, and do not necessarily reflect the true nature of reality, what basis do we have to assume that this ‘social skills and sympathies’ narrative will help us make more accurate predictions within the domain of science of mind than a sound and rigorous neurobiological model? Although the latter is indeed but a domain specific approximation as well, we have good reason to assume that it makes accurate predictions by drawing correlations between mental events and neural events. Like every other scientific domain, there are still ‘correct’ and ‘incorrect’ models within our perspective based paradigm that we use. The level of predictive accuracy of a certain model is the litmus test for its legitimacy and usefulness. Even though Newtonian mechanics is an approximation that breaks down past the quantum scale, Newtonian laws provide far better predictive power than Aristotelian physics – thus the former is considered to be the 'correct' model. In this case, various proposed models of consciousness, such as IIT or 'Global Workspace' could still be 'wrong' relative to each other depending on which one is able to make more accurate predictions. Some approximations are clearly much more effective than others, and using 'social skills' is hardly a rigorous way to construct a predictive model 'narrative'. Although we can never develop a truly 'first person' predictive model 'narrative' because we cannot access other first person perspectives, we can construct rigorous third person models that make more accurate predictions.

What would a 'correct' Stage Two scientific model look like? Let us assume IIT is the correct scientific model of consciousness as an example. Within the IIT model, the physical things we experience as much grander than ourselves, mountains, rivers, stars, galaxies, are actually miniscule when it comes to integrated information (Tononi 2008). We simply represent these things grandly within the physical way of viewing reality. When it comes to integrated information way of viewing reality however, our brains with which a high level of phi makes them greater than any of the aforementioned. Tononi concocts a hypothetical device called a 'qualiascope', which can allow us to represent reality in terms of integrated information, rather than physical entities. With the qualiascope, mountains and galaxies will be represented as miniscule, negligible qualia shapes, but human (and other vertebrate) brains will be represented as grand constellations of complex qualia shapes. The framework established around integrated information theory may provide us with 'laws' that can accurately predict behavior within the specific domain of science of mind, but that do not necessarily carry over to other domains, like how Newtonian laws make accurate predictions within classical mechanics but do not carry over to quantum mechanics. This new way of viewing reality may give a much more accurate predictive model within science of mind. In this case the IIT model would serve as the 'correct' model of consciousness.

It is important to keep in mind that this is still a Stage Two model; the qualiascope will not allow us to observe the underlying neutral reality, it simply another third person objective representation of it. We will see grand constellations of complex qualia shapes with integrated information, but will not be able to translate between integrated information qualia shapes and subjective experiences. The 'qualiascope' gives us but an alternate way to represent reality objectively, as 'shapes of integrated information' rather than 'physical matter', to allow us to make more accurate predictions within a specific domain. Although it will not allow us to bridge the explanatory gap and translate between the first and third person perspectives, this model will still serve as a practical and effective science of mind in that it, like good models in any other domain of science, provides highly accurate predictions which offer a variety of different applications.

10. Proposed Experiments to Help Us Reach Stage Two

The following is a description of proposed experiments to help us achieve Stage Two in the process of building a science of mind. They are not intended to be a comprehensive outline of all the necessary experiments needed to reach Stage Two, as there are potentially an infinite number of experimental setups that could aid in establishing specific neural correlates of specific qualitative aspects of consciousness. These are simply representative of the type of experiments that need to be conducted in order to accomplish a Stage Two framework. It is important to keep in mind that this is still geared toward establishing a framework based on correlations. A comprehensive and highly accurate predictive model based on correlations will serve as a launch-pad to possibly go beyond Stage Two, but whether Stage Three is possible remains yet to be seen. Even if Stage Three is not reachable, Stage Two can nevertheless constitute an effective and practical science of mind.

An ideal experiment to help bypass the problem of the change in perspectives would be to have experimental neuroscientists observe their own 'physical' brain states on a monitor, while at the same time noting their own corresponding mental experiences. Since an experimenter can never experience what the subject is experiencing and can only rely on first person reports, and since first person reports are often misleading, it would be helpful to eliminate the middle-man and have the experimenter observe his own brain states while at the same time noting his own subjective experience. To replicate the results, other experimenters would need to agree on the specific subjective experiences everyone has correlating to specific induced brain states. This type of study is not technologically possible as of yet.

Visual awareness has thus far been studied intensively, but what about other senses, such as auditory awareness, for example? Perhaps NCC studies can be run on subjects who have been blind from birth (congenital blindness) or lost sight later in life (later-onset blindness), studying the differences between the two as well as from control subjects with normal vision, and if certain regions of the visual cortex exhibit activity associated with sight in blind patients, when they hear certain sounds or use sound to echolocate. It would be interesting especially to study individuals who have later-onset blindness to see the activity of the visual cortex during dream states when the subject can still 'see' visual images. For example, do the neural correlates of a blind subject's experience of a specific color in a dream compare similarly to the neural correlates of a non-blind subject's experience of the same color in a waking state? Some studies have even suggested congenitally blind individuals are able to have 'visual' awareness during dreaming. An EEG study found that alpha wave activity (8-12Hz) becomes markedly stronger in the visual cortex of not just non-blind, *but also congenitally blind*, subjects, during periods when the subjects report "visual" content in their dreams (Bértolo et al. 2003). Depending on how similar their brain activity – specifically in the visual cortex – during dreaming is compared to that of non-blind individuals during dreaming, we could assess how similar congenitally blind individuals' 'visual' dream experiences are compared to visual dream experiences of non-blind individuals. All these experiments would be especially useful to pin down neural correlates of very specific qualitative aspects in visual awareness.

These experiments can also be applied to subjects who are not visually impaired. Experiments have already been conducted that have identified general differences in brain activity between dreaming and waking states (Hobson et al. 2000). A recent study has also attempted to reconstruct the visual experiences of subjects during dreaming, based only on the subjects' brain activity (Horikawa et al. 2013). While the subjects were sleeping, their brain activity was monitored using fMRI. The subjects were awakened periodically and asked to describe the visual experiences in their dreams. These descriptions were organized and correlated with the monitored brain activity, then developed into an algorithm which could 'translate' between general categories of dream imagery and types of brain activity. This algorithm was then used to successfully predict, with around 60% accuracy, general categories of dream imagery from the subjects' brain activity during dreaming. This study found that visual dream imagery involved much of the same type of visual cortical activity associated with 'real' visual imagery (corresponding to actual external stimuli) during waking state. General correlations between visual cortical activity and qualitative visual experiences in dreams have been identified, but the next step would be to test the neural correlates of *specific* qualitative aspects of visual or other sensory experiences during both dreaming and waking states. A subject's brain activity can be monitored for neural correlates during dream states involving very specific qualitative aspects. These neural correlates can then be compared for similarity to the neural correlates associated with the experience of the same qualitative aspects during waking state. For example, if the subject smelled pot roast or saw the color blue in a dream during which his neural correlates were monitored, these correlates can then be compared to correlates corresponding to the subject's olfactory perception of a real pot roast or visual perception of the color blue during waking state.

Certain aspects of our experience correspond unequivocally to an external reality, but how we interpret these aspects varies from individual to individual. We should be able to test for neural differences between someone who is experiencing something for the first time and someone experienced in that specific area. For example, a subject untrained in facial recognition of chimpanzees can study a collage of individual chimpanzee faces while his neural activity is monitored. After training him on facial recognition of chimpanzees, have him study the same collage of faces. This experimental model can be applied to other senses as well. In the case of auditory sense for example, have a subject completely ignorant of Bach listen to the Well-Tempered Clavier, while monitoring his neural activity. After training the subject to recognize the specific musical patterns and structures in Bach's music, let him listen to the same recording of WTC on the same equipment in the same environment. Individuals who are trained and experienced in these respected fields (e.g. 'Bach experts') can be used as control subjects. Will we be able to detect neural changes? If yes, this implies that although *what is being represented* in terms of the input, 'objective third person' has not changed, the *representation*, 'subjective first person' has. What areas of the brain, the visual cortex, are associated with this change, specifically? Which parts of the visual cortex, in the case of the visual awareness tests, remain identical in activity before and after the training, and which parts exhibit changed activity? These experiments will help us to make a better distinction of what, in neurobiological terms, is considered 'objective' as opposed to 'subjective'.

There are many potential experiments one could design to study neural correlates in the visual cortex. For example, in change blindness, what difference in visual cortex activity, specifically in the higher visual centers, is there before individuals notice a change in an image, and after? One would assume that their visual input is identical in both cases, since the image remains identical before and after a change is noticed. Based on the results in the binocular rivalry experiments, this would indicate that change in V1 activity would not be significant. Another experiment would be to test visual awareness with well known optical illusions, such as the duck-rabbit illusion. Although the visual input into the primary visual cortex remains identical, the subject only experiences either a duck or a rabbit at any given time. Studying the neural differences in the extrastriate visual cortex areas between when the subject experiences 'duck' or 'rabbit' would aid in determining the neural regions responsible for conscious visual representation, as opposed to mere integration of visual stimuli. The tests on macaque monkeys involving binocular rivalry are similar in nature to these proposed experiments, and it would be interesting to compare the visual cortical patterns of activity in these experiments to those in the binocular rivalry experiments.

Studying neural correlates of specific color experiences can give us insight into the nature of qualitative aspects of our representations of the 'non-self' external world. Establishing the neural correlates of specific colors is relatively straightforward, because we can use color illusions to narrow down the specific brain activity associated with experiencing a specific color, regardless of whether the experience corresponds to the actual color in the external environment. The phenomenon of negative afterimage, seeing an inverted color after the original color stimulus has disappeared, is a well known and studied color illusion (Shimojo et al. 2001). An example of negative afterimage would be that a subject reports seeing the color green after staring at the color magenta (green's inverse color) for an extended period of time, followed by a quick transition to a white background. Negative afterimage occurs because the ocular photoreceptors corresponding to a particular color, e.g. magenta, become inactive when over-stimulated, and begin sending weaker signals to the brain. When the magenta stimulus suddenly disappears, the over-stimulated photoreceptors continue to send out weak signals, but the surrounding photoreceptors not corresponding to magenta have not been over-stimulated and send a strong signal in comparison, identical to the signal sent when they are actually being stimulated by the color green, the inverse of magenta. In this case, photoreceptors send the same 'green' signals to the brain regardless of whether it is a response to an actual green stimulus in the external environment or whether it is a response to the sudden disappearance of the inverse magenta stimulus. Thus all neural correlates corresponding to 'seeing the color green' should be identical, whether in the case of experiencing an illusion of green, or the case of experiencing a color green actually present in the external environment. This implies that specific qualitative aspects are coded for by specific neural activity, i.e. the qualitative aspects of our subjective experience are entirely a product of the integrated system 'self' – the brain. 'Qualia', or what things appear to be like, has no bearing on the actual nature of the external reality. This of course makes sense as qualia only serves to aid organisms in distinguishing the differences and similarities between external objects as a form of reference.

Certain optical illusion images are able to give the viewer the false perception that the picture is moving, while it is in fact stationary. The V3 area of the visual cortex is known to be correlated with perception of dynamical motion, or the feeling of movement within a certain spot, such as rotation or undulation (Zeki et al. 2003). It would be interesting to monitor the activity of V3 when a subject is viewing the 'moving' optical illusion images, versus when he is viewing a non-illusory stationary image, versus when he is viewing an actual dynamical moving image. If V3 activity is found to be similar with respect to both the illusory 'moving' images and the actual moving images, and different with respect to a non-illusory stationary image, it would further strengthen the idea that qualitative aspects of our representations are internal constructions that do not necessarily reflect the properties of the external neutral reality.

Additionally, it would be interesting to investigate the similarities of the neural correlates associated with the *imagination of experiencing* a particular qualitative aspect to the neural correlates associated with the *actual experience* of that qualitative aspect. For example is the brain activity associated with imagining a sharp pain in the left foot similar to the brain activity associated with actually experiencing a sharp pain in the left foot? These types of experiments could provide us valuable insight on what actually must occur in our brains for us to truly understand what it is like to have a particular experience. An important application of this is that it may help us to determine whether a Stage Three understanding within the science of mind is truly reachable. For example, in order to independently deduce the qualitative aspect of a mental event from a novel brain event whilst having no previous knowledge of experiences of this type (e.g. a colorblind scientist trying to understand the qualitative aspects of various colors), as is required by a Stage Three understanding, one would somehow have to activate regions of the brain responsible for producing¶ these experiences, or at least the imaginations of these experiences. The question is, how similar is the brain activity responsible for producing the *imagination of experiencing* say, the color red, to the brain activity responsible for producing the *actual experience* of the color red? Based on an assessment of these similarities we could objectively determine whether an imagination of the experience would entail a true understanding of what the experience itself is like. This would help us get a better picture of what is necessary for a Stage Three framework to accomplish (i.e. what types of neural activity can be induced within our own brains from an application of the framework, whether the neural activity that produces an imagination of the qualitative aspect would be adequate, based on how similar it is to the actual experience itself), in order to constitute a true understanding of the science of mind.

Finally, recent advancements in computational neuroscience have actually allowed us to roughly reconstruct visual experiences using computational models of the visual system (Nishimoto et al. 2011). To create the computational model, the brain activity of subjects is monitored using fMRI while they watch many hours of video clips. The computers then construct an algorithm that correlates specific qualitative aspects in the video clips, e.g. specific shapes, colors, motion, with specific types of measured brain activity in the subjects while they were watching the clips. This algorithm can then be used to predict *novel* qualitative experiences the subjects are having, by only observing their brain activity using fMRI (Fig. 10.1). For example, if the subject is watching a video clip of a woman's face, fMRI measures the brain activity of the subject, and the computer algorithm then 'translates' it to its correlated specific qualitative aspects, i.e. the woman's hair, nose, eyes, mouth, which is organized to reconstruct a rough image of what the subject is actually seeing. Although still in its preliminary form, this is the beginning of a Stage Two predictive model, in which an algorithm, based on a highly accurate correlation between specific qualitative aspects and specific neural activity, can predict the mental events corresponding to novel brain events. It is important to keep in mind that this is still an *artificial algorithm based on neural correlates*, that is, it does not provide a true explanation of how and why a given brain event translates to a given mental event. Nevertheless, it is a wonderful and elegant model of brain-mind correlation that is consistent with the RPNM paradigm –that specific neural activity in the brain does in fact correspond to specific qualitative aspects in a subjective experience. However, whether or not they are the third and first person representations of the same neutral reality, respectively, can only be determined by a successful Stage Three framework.

¶I am using physicalist jargon here for the sake of simplicity. The brain events do not actually 'produce' the mental events; 'mental' and 'brain' events are the first and third person representations of the same neutrality, respectively.

11. Perhaps Stage Three Can Be Reached After All, but Doing So Will Require Expanding Our Scientific Paradigm

Ultimately, it is tempting to say neuroscientific models can 'explain' consciousness insofar as scientific models can 'explain' anything, because everything is really just an approximation used to make predictions, and just leave it at that. The approximate nature of scientific laws is just more obvious in science of mind because of the change in perspectives from third to first person. However, this is an incredibly unsatisfying conclusion and could even be seen as yet another deliberate attempt to avoid solving the real problem of consciousness. Perhaps Stage Three can be achieved after all, and explanatory bridge laws that provide the link between the mental and the physical can be established. This will need to happen for us to truly understand the nature of the neutral reality, and the first and third person representations. Even if Stage Three is ultimately unattainable, we will never know unless we pursue it.

When will we have an explanatory science of subjective experience? This will occur when we have established the explanatory relationship between neuronal mechanisms and certain experiences. How will we know when this has happened? As David Chalmers noted, there are certain phenomenal truths (such as the redness of red or timbre of a clarinet) which cannot be deduced but can only be known through experience itself (Chalmers 2013). We could conceivably understand exactly how the brain 'simplifies' a neural firing pattern into a specific experience but a colorblind or deaf person would still not be able to use that understanding to deduce what 'redness' or the timbre of a clarinet was. Until this can happen, the science fails. Because the subjective aspect of the experience is itself the end goal, unless you can deduce 'redness' from the knowledge of brain processes, you really *don't* understand how the brain simplifies these processes into 'redness'." So we will know explanatory relationships have bridged the mental and physical when colorblind Mary will be able to infer exactly what the subjective experience of seeing red would be like purely by understanding the neural events which translate into it. Or when we ourselves would be able to predict what a novel experience would be like subjectively/qualitatively in every possible pattern of neural events. Our comprehensive framework of classical mechanics for example, is able to predict based on Newtonian laws exactly what results (outputs) will occur given a certain series of every possible set of preconditions (inputs), with 100% accuracy. A complete science of subjective experience thus, must be able to predict based on our comprehensive understanding of the fundamental translative bridge laws, exactly what subjective experiences will result given a certain pattern of neuronal activity, with 100% accuracy.

As noted above, identifying and isolating very specific neural events associated with very specific qualitative aspects to predict novel qualitative experiences, e.g. event X associated with seeing green color, event Y associated with seeing cube shapes, ergo X and Y simultaneously result in experience of a green cube, still only gets us to Stage Two. Individuals lacking knowledge of any qualitative aspects will not be able to predict these novel experiences given the neural events. Even if a colorblind person identified the specific neural event associated with seeing green color, in the above example, he would still not be able to infer from that information alone what it would be like to experience the 'greenness' of the cube. To use another analogy, suppose someone without knowledge of Newton's laws of motion was able to identify that hitting a bowling ball with a plank is associated with the ball moving, and that a ball hitting a rubber wall is associated with it bouncing backwards. He would have identified some specific correlations with which he could predict some novel events given that they were associated with the known correlations, e.g. hitting a ball with a plank towards a rubber wall causes the ball to move forward, hit the rubber wall, and bounce backwards, but he would still not understand *how and why* this was true. Furthermore, he would not be capable of predicting the result of a completely novel set of inputs governed by the same laws, e.g. the behavior of a rocket fired with an internal combustion engine, without having observed the result first. If he understood Newton's laws however, he would be able to predict with 100% accuracy the results of every possible series of inputs, without having knowledge of any correlations or witnessed the result beforehand (how else would we have reached the moon?).

Thus to truly formulate a complete objective science of subjective experience, we must understand the *translative* relationship between neural firing patterns and subjective experience, the "laws of consciousness," if you will. These "laws" must be able to explain *how and why* a particular neural event translates to the mental event/subjective experience of seeing red. To successfully accomplish this would mean that if we induced a novel pattern of neuronal firings never before carried out on any living organism, we would be able to predict, based solely from our understanding of these “laws”, exactly what the subjective experience coupled with it would be like, without having actually experienced it ourselves beforehand. A deaf person who understands these “laws” would be able to, after studying the neural mechanisms associated with hearing the sound of a tuba, know exactly what the subjective experience of hearing a tuba would be like – without having ever heard a tuba. And if Mary understood these “laws” she would be able to know exactly what subjectively seeing red would be like, even if she never gained the ability to distinguish color. This ambitious goal is unattainable within a neuroscientific framework based purely on establishing neural correlates of consciousness. However, if we can ever achieve a thorough conceptualization of the neutral reality from which both first and third person perspectives are derived, we could potentially understand how and why one translates to the other.

As I mentioned above, the nature of our perception entails that we cannot directly observe or measure the neutral reality, so how would it be possible to establish translative bridge laws between the mental and physical? RPNM states that the true reality underlies our current conception of the universe, based within the context of our self-external relationship maps. We might not be able to directly ‘access’, that is, observe, the ‘neutral’ reality, but RPNM certainly does not necessarily entail that it is impossible in principle to develop a conceptual understanding of that reality. From this understanding of the ‘neutral’ reality, we will ideally be able to deduce both the third and first person perspectives. These are not laws of causation between the mental and physical, as mental is not caused or generated by the physical, but rather explanatory bridge laws from which we could derive both the third and first person perspective from the same neutral reality. Once that is achieved, a true science of mind will have been completed, one that can predict novel mental events given novel neural events, and vice versa, not because of a highly accurate artificial translative algorithm, but because of a true understanding of the underlying bridge laws that translate between the two perspectives.

What would a Stage Three framework actually have to manage to accomplish to achieve this explanatory power? Based on the 'mental'-'physical'-neutral translative relationship as summarized in Fig. 7.2, a rough sketch of a Stage Three understanding is outlined as follows. A subject experiences a 'mental event', which an observer sees as a 'physical brain event' in the subject. The observer wants to translate his perception of the 'physical brain event' into the subject's perception of the 'mental event'. He knows that this 'brain event' is his third person representation of a neutrality 'A' occurring in the subject's brain. However, this 'third person' representation also constitutes *the first person representation of a separate neutrality 'B' occurring in the observer's own brain*. The observer's third person representation of 'A' is identical to his first person representation of 'B', so translating his third person representation of 'A' into the subject's first person representation of 'A' would be the same as translating his first person representation of 'B' into the subject's first person representation of 'A'. How does the observer go about doing this? First, must use the first person translative bridge laws to translate his first person representation of 'B', the 'brain event' – also his 'third person representation' of 'A', into the neutrality 'B'. Then he must use the third person translative bridge laws to translate 'B' into its corresponding 'A', based on the knowledge that neutrality 'B' is activated in his own brain when he observes neutrality 'A' in the subject's brain. Finally, he must use first person laws again to translate neutrality 'A' into the subject's first person representation of 'A', the 'mental event'. In order to deduce from this process what the subject is experiencing however, the observer must, through application of the translative bridge laws, somehow activate his *own* neutrality 'A' so that he can also, like the subject, experience the first person representation of 'A', the 'mental event', or at least be able to imagine what the first person representation of 'A' is qualitatively like. This entails that if the observer was a colorblind scientist for example, a true understanding of the translative bridge laws would actually allow him to activate neutralities in his own brain never activated before: those represented as color experiences in the first person perspective. In order for a Stage Three framework to allow individuals having total lack of known experiences to be able to nevertheless predict the qualitative nature of an experience corresponding to a novel neural event, it must be able to accomplish the aforementioned.[‖]

However, contemporary neuroscientific approaches are geared toward establishing neural correlates and descriptive models of consciousness. If we continue down this path, we can attain Stage Two, but will not progress any further. Some may consider this an adequate enough scientific framework, but it ultimately does not satisfy our thirst for a true understanding of how to bridge the mental and physical. Our current paradigm in neuroscience does not gravitate toward models that will allow us to conceptualize the underlying neutral reality. A complete understanding of reality must be all inclusive, and since our current 'physicalist' paradigm obviously does not encompass the mental, a complete understanding of consciousness would require us to expand our paradigm or a paradigm-shift. Instead of narrowing focus on the most intuitive approaches to the problem, such as establishing specific neural correlates, it is important to incorporate other scientific domains in order to get the big picture. Our current focus only on mainstream neuroscience is analogous to a scientist breaking a radio apart to find the musicians playing inside, whereas in reality the source of the sound comes not from within the radio itself but from transmitted electromagnetic waves, or scientists' original approach to studying the phenomenon of light by attempting to find its 'essence', before Maxwell's equations demonstrated visible light was simply a form of electromagnetic radiation. If we expand our inquiry into a broader scientific framework, we begin to see a pattern emerging. Many domains of science have reached a seemingly impassible barrier within the past number of decades, in which much evidence seemingly contradicts our understanding of reality. When considered as a whole, this comprehensive, but not conclusive, evidence from a large number of scientific domains, largely ignored thus far, begin to paint the picture of a reality that is not grounded in our 'physicalist' model. With this comes the possibility that perhaps consciousness is ultimately based not neurally, but on something more fundamental. Much of this evidence has been dismissed out of hand due to none other than the fact that it does not fit in with our current paradigm of reality, the 'physicalist' assumption. Conversely however, it is also important not to be too hasty; the interpretations and conclusions drawn from this evidence are at the moment still incredibly speculative. Nevertheless, the evidence is stronger than most would give it credit for, and is at least worth serious consideration, as current scientific progress seems to be static in a number of different domains.

||One may argue that although a Stage Three framework may allow us to understand how to translate between the mental and physical, it is possible we still cannot use it to independently deduce mental events corresponding to novel brain events. So a colorblind scientist could potentially understand the Stage Three framework without being able to use it to deduce the qualitative nature of color experiences. It is true that even if one is incapable of using a Stage Three theory of mind to deduce mental events from novel brain events, it does not invalidate the theory. However, one cannot *truly understand* the Stage Three theory unless he is able to use it to independently deduce mental events from novel brain events. In any other domain of science, understanding a theory will entail one to deduce all aspects of the phenomenon the theory is describing, independent of any preexisting knowledge of the phenomenon. For example, a complete understanding of the photosynthetic pathways will allow one to

independently deduce carbohydrate biosynthesis without having any preexisting knowledge that carbohydrate synthesis is involved in the pathway. If one cannot deduce carbohydrate synthesis from his understanding of photosynthesis without having someone else make the statement claim 'photosynthesis includes carbohydrate synthesis', it means he has not really understood the phenomenon. Likewise, a true understanding of a Stage Three theory of the mind should allow one to independently deduce the first person mental from the third person physical without having to make a statement claim of 'this process describes the translation of the physical to the mental' at any point. As I noted above, if this type of deduction is not achievable it will not have invalidated the Stage Three theory, as this process may still be the accurate description of the translation between the first and third person perspectives. However, if we are not able to achieve this type of deduction, it will mean we have not truly understood the theory. Therefore, if application of the Stage Three framework does not allow us to activate neutralities in our own brains never activated before in order to deduce mental events from novel brain events, it will mean we are neurobiologically incapable of truly understanding the Stage Three theory. Although this may pose a potential problem in the future, at the moment it is better to put this concern aside and focus on establishing a Stage Three theory in the first place.

12. How Could Quantum Physics Help Us Reach Stage Three?

Why would we have any reason to look beyond neural activity and correlates to something more fundamental in order to understand the link between mental and physical? Quantum hypotheses of the mind, although supported by a number of prominent physicists, have been largely dismissed by the mainstream neuroscience community. Is there a deeper connection between quantum mechanics and consciousness other than 'Both are mysterious and therefore must be related', or a passing New Age fad?

The double slit experiment demonstrates a very curious link between measurement/observation, or as some physicists would argue, conscious observation, and the way reality behaves (Weeks 2013). A classical double slit experiment is conducted with shining light waves through double slits onto a receptor plate. Because the light is going through two slits, it comes out as two waves which interfere with each other, producing an interference pattern on the plate. Conversely, if small particles (such as beads) are fired through the double slits, they produce a two line pattern upon hitting the plate, corresponding with the double slits. The original experiment that demonstrated wave-particle duality was conducted by firing electrons through double slits onto a photographic receptor plate (Fig. 12.1). Doing so resulted in an interference pattern appearing on the receptor plate, as if the electrons were behaving like waves going through both slits at the same time, rather than single particles either going through one slit or the other. It was initially believed that the pattern was a product of the electrons bumping into each other, but firing single electrons through the slits one at a time produced the same interference pattern. This was strange enough, but the result of the subsequent experiment was even stranger. To get to the bottom of this mystery, physicists placed a measuring device beside the double slits on the side the electrons were going in, to detect which specific slit each electron went through. The rationale was that by doing this one could be able to observe why the electrons were behaving as if they went through both slits at the same time. However, as soon as a measuring device was placed beside the double slits, the electrons' behavior changed; this time they behaved as particles; the measuring device observed discrete particulate electrons going through either one slit or the other. Furthermore, the pattern on the receptor plate changed after a measuring device was introduced; instead of an interference pattern, the electrons left a double slit pattern! So if no measuring device is present to observe electron behavior going through the double slits, electrons behave like waves, and if a measuring device is present, the electrons behave like particles. This became known as the wave-particle duality paradox. The results do not fit anywhere within the traditional physicalist model of the universe yet have been repeated with astounding accuracy in almost a century of continuous variations of this experiment.

The Copenhagen Interpretation of the wave-particle duality phenomena provides a mathematical model with tremendously accurate predictive power, but is ultimately, like simply establishing specific neural correlates for a science of mind, unsatisfying, as it does not seek to understand what the evidence means: how is matter behaving in two different ways, depending on whether observation occurs or not (or how is neural event 'A' translated into mental event 'A')? A central theme within the Copenhagen Interpretation of wave-particle duality is complimentarity, which is taken to be an integral property of quantum phenomena, just how it is. In short, complimentarity states that the quantum phenomena, depending on the type of measuring device used, behaves either like a particle or a wave, which can both be described in classical terms (Walker 2000). According to the complimentarity principle, a measuring device which can observe both the wave and particle aspects concurrently is logically impossible. Thus it is useless to attempt to achieve an explanatory bridge between the two aspects of wave-particle duality. One cannot help but feel this is analogous to the aforementioned idea that since it is logically impossible to observe the third and first person perspectives simultaneously, one should not attempt to bridge the gap between the mental and physical, and focus instead on establishing a model based on drawing correlations between mental and neural events. Einstein famously expressed his opinion that this was a dismissive way of addressing the dichotomy, and suggested the paradox, rather than just being the way things were, hinted at an incomplete paradigm within physics (Bohr 1958).

As I hinted at above, wave-particle duality and the complimentarity principle of two simultaneously unobservable aspects of the same phenomena is analogous to the seemingly 'unbridgeable' duality of the first and third person perspective. Perhaps there is something more to this than a mere analogy? Depending on whether the experiment is measured/observed, matter exhibits wavelike or particle-like behavior. A lack of observation is equivalent to a true third person ‘objective’ approach, whereas active observation involves the first person perspective. In other words, here reality is being approached from two different points of view, much like in the study of consciousness as either neural (third person) or mental (first person) events. However, instead of making identity claims between neural and mental or declaring complimentarity, two aspects of the same phenomena, as an inherent feature of matter, RPNM opens another possibility: an underlying ‘neutral’ reality from which both aspects can be derived.

Many physicists, unhappy with the Copenhagen Interpretation, recognized a curious link between consciousness and quantum theory. They noticed that the process of measurement/observation is a central and imperative theme in quantum theory. John von Neumann described the process in mathematical terms as a “discontinuous, non-causal, instantaneous and irreversible act given by the transition of a quantum state to an eigenstate (measured state)” (Quantum Approaches to Consciousness 2013). This is commonly known as the ‘superposition’ and subsequent ‘collapse of the wavefunction’. In other words, the collapse of the wavefunction entails a reduction of possible states of the electron, called a superposition, into a single ‘collapsed’ state (Weeks 2013). For example, in the double slit experiment, the electron is in a superposition of many different possible states/locations. This superposition of probabilities behaves as a wave, which, because it travels through double slits, becomes two ‘probability waves’ that interfere with each other. However, these superposed states collapse the moment they reach the photographic receptor plate, as that is the place where the electron’s location is measured. The reason the wavelike interference pattern appears on the plate is because up until the moment the electron reaches the plate, it still behaves as two superposition waves (by way of the one wave traveling through the double slit) of all possible locations, interfering with each other. When the electron hits the plate, the measurement occurs and the superposition collapses, the electron now being in a discrete location anywhere within the former superposition wave. After many electrons hit the plate in this manner, the interference pattern emerges. However, the electron is observed by a measuring device *before* it enters the double slits, the superposition of possible states/locations of the electron 'collapses' into a single state/location, such that it becomes a particle in a discrete location going through either one or the other slit. In this case no interference pattern emerges because the electron left the slits as a particle, not as a superposition wave. This carries two implications. First, the wavefunction collapses as soon as measurement takes place, and second, the wavefunction collapse occurs in a distinct location in the universe – in the case of the double slit experiment, it appeared to occur where the electrons were measured by the photographic receptor plate (interference pattern) or where they were measured by the device before entering the double slits (no interference pattern). Several physicists drew the conclusion from this that consciousness was the necessary agent to collapse the wavefunction, and that in fact the only real location of wavefunction collapse was in the conscious brain of the observer. Fritz London and Edmond Bauer were the first to postulate that in fact conscious observation caused the wavefunction collapse, completing the quantum measurement (London and Bauer 1939). Their reasoning was that even when a superposition is allegedly collapsed by a non-conscious measuring instrument, the result cannot truly be observed and measured until a conscious observer is present. Eugene Wigner's thought experiment 'Wigner's Friend' further illustrated this concept, attempting to show how consciousness was necessary for quantum measurement (d'Espagnat 2005). The thought experiment runs as follows: Wigner's friend conducts Schrodinger's Cat experiment when Wigner is not present in the lab. When Wigner returns he learns from his friend that the cat is alive/dead. So where does the wavefunction collapse occur? On the cat, when the friend opens the box to see if it is dead or alive? On the friend, when Wigner enters the room to see if he is happy or sad? Without Wigner coming in to observe his friend, would his friend be in a superposition of happy/sad, as well as the cat being in a superposition of alive/dead? It becomes an infinite regression of superpositions *until* – the conscious observer or measuring device becomes involved. Wigner deduced from this thought experiment that collapse must be relative to the measuring observer. The friend collapsed the wavefunction with respect to the friend when he opened the box; Wigner collapsed the wavefunction with respect to Wigner when he entered the room – the collapse occurs in the mind of the conscious observer. After initially postulating that the observing system could be either a human brain or a non-conscious measuring instrument, von Neumann in his later career concluded in agreement with London, Bauer, and Wigner, after struggling to find where exactly the wavefunction collapses in the process of measurement. Attempting to fit the location of the wavefunction collapse within the external universe, von Neumann introduced the concept of the 'von Neumann chain', in which the measurement process was broken down into an infinite number of steps. Anywhere along this chain a non-conscious measuring device could potentially collapse superpositions into single states, but if a measurement and subsequent collapse were to occur anywhere along this chain, the universe would be split into a superposed and a collapsed state, which was logically incoherent. The *only* place that made sense to insert the collapse was where the final measurement took place: the conscious mind of the human observing the experiment. Thus he concluded that the site of wavefunction collapse was ultimately located within the brain of the conscious observer.

An additional reason to draw a link between quantum theory and consciousness involves quantum indeterminacy and free will. According to Henry Stapp, the wavefunction collapses as a result of the brain selecting a quantum possibility among many alternatives (Quantum Approaches to Consciousness 2013). This opens up the door for free will. Even philosopher John Searle, who was originally callously critical of quantum models of mind, acknowledged in the last several years that the only example of indeterminate processes in nature lay within quantum theory, and that quantum mechanics was the only currently conceivable way to salvage our intuition of free will without explaining it away as an illusion (Searle 2007). Naturally, indeterminacy itself does not equate free will, but as Stapp proposed, the possibility is selected for among a number of indeterminate states.

A third reason to draw a link between consciousness and quantum physics is the phenomenon of quantum entanglement, what Einstein called "spooky action at a distance". Quantum entanglement is best illustrated with the Einstein-Podolsky-Rosen (EPR) Paradox and Bell's Theorem (Fig. 12.2). In the EPR Paradox, an unstable particle with a quantum spin 0 decays into two particles, A and B, which each must have a spin of either -1/2 or +1/2, to add up to the 0 of the initial particle (Felder, 1999). However, quantum physics states that rather than having a definite spin of either -1/2 or +1/2, both particles A and B exist in a superposition of both -1/2 and +1/2. When one particle is observed/measured however, it collapses into a single spin state of either -1/2 or +1/2. But because both particles are decayed from the same parent particle, if A has a -1/2 spin, then it must follow that B has a +1/2 spin, and vice versa. Therefore, according to quantum physics, as soon as particle A is measured, particle B must instantaneously collapse into a single spin state as well. But what if the decayed particles are separated by a great distance? An immediate collapse of B to a single state following a measurement of A would imply an instantaneous, non-local transfer of information. Bell's Theorem postulated that if faster than light speed information transfer was impossible, particle B would have to be in a superposed spin state even after particle A was measured. Thus, if B is subsequently measured before the information of A's specific collapsed state can reach it, based on normal probabilities, one would have to expect observing either a -1/2 or +1/2 spin state on B, independent of A's spin state. However, numerous experiments have shown, with increasing distance up to several kilometers, that no matter how far the particles are separated, as soon as A is measured and its state collapsed, B instantaneously collapses to the opposite spin state. This not only contradicts Newton's Laws of motion, but also indicates faster than light speed information transfer. How does this relate to consciousness? According to classical models of neural communication, there must be a time lag between the activation of one neuron and it's signaling via action potentials to neighboring neurons. However, the phenomenon of 40Hz gamma oscillation, the most well demonstrated brain-wide neural correlate of consciousness, is an instantaneous communication, meaning that there is no time lag between when neurons in separate and opposite areas of the brain are activated. When gamma oscillation occurs, all neurons within the brain fire at the same time in perfect synchrony.

How would a quantum approach to the problem of consciousness accomplish Stage Three within RPNM? Most intriguingly it specifically addresses an issue which may have been plaguing the reader up to this point. Establishing translative bridge laws between the first and third person perspectives rests on the assumption that there *are* in fact such laws, which translate between specific qualitative aspects of the first and third person representations and the properties of the neutral reality. This would mean that rather than the qualitative aspects of our representations being an arbitrary product of evolution to allow us to distinguish various elements of the external environment from each other, there are in fact intrinsic relations between what representations 'are like' and the properties of neutral reality, which evolution has facilitated to allow biological organisms to utilize. For example, the type of qualitative aspect of the color red is an intrinsic property that is translated from a particular type of neutral property. This is analogous to a type-type relation – the aforementioned however is applied within physicalist identity theory between mental and brain states – in which every type of mental event is identical with a corresponding specific type of brain event (The Mind/Brain Identity Theory 2013). In type-type identity, the type of brain event corresponding to my experience of red, when occurring in someone else's brain, would correspond to his experience of red as well. Conversely, in a token-token relation, although mental events and physical events are identical, every type of mental event does not correspond to a specific type of brain event. In token-token identity then, the type of brain event corresponding to my experience of red, when occurring in someone else's brain, would not necessarily correspond to his experience of red as well. If the relation between the neutral reality and our representations of it were analogous to token-token identity, no bridge laws would exist to translate between neutral reality and our representations of it. To illustrate this concept further, imagine a hypothetical extraterrestrial species. On this planet, even if the qualitative aspects of our representations are arbitrary products of evolution, we have reason to believe the qualitative experiences of our representations of reality are at least similar if not identical, as all animals on this planet share a common ancestor. However, this alien species has evolved independently of life on earth, so would the qualitative experiences of their representations of reality be similar to ours? For example, would they experience the neutrality we humans experience as red color the same way we do? If qualitative aspects of representations are an arbitrary product of evolution, then the answer would almost certainly be no. However, if translative laws intrinsic to reality govern between specific types of neutral reality properties and specific types of qualitative aspects, the answer would be yes. In this case, Stage Three could be achieved by understanding these laws, such that say for example a colorblind neuroscientist, following these laws, could deduce the specific qualitative aspect (first person perspective) of red color based on its corresponding neutral property (which was derived from the third person perspective of it). The predictive model established in Stage Two will never definitively prove that these translative laws exist, as it is merely a correlative framework. However, quantum theory not only implies the laws do exist, but actually provides a possible way in which we could understand them and the neutral reality.

Going beyond the mathematics of the Copenhagen Interpretation and understanding the metaphysical implications of the wavefunction collapse may allow us to in fact conceptualize the nature of the neutral reality. Having understood the neutral reality, we could then possibly provide the explanatory law(s) that will bridge the neutral reality and the first and third person representations of it, 'mental' and 'physical', and consequently be able to translate between the two perceptions/perspectives of the neutral reality. Confirming this would also demonstrate that the qualitative aspects of our representations are not arbitrary, but that there are explanatory laws within reality which evolution has facilitated, that translate between neutral reality and the qualitative aspects of its corresponding representations. But what reason do we have to believe the wavefunction collapse can provide the translative laws between neutral reality and the first and third person representations? The wavefunction collapse is inexplicably linked with the act of observation and representation. Ultimately, something happens at the moment of observation, when the superpositions collapse to single states. As von Neumann argued, the only place a true wavefunction collapse could occur within the conscious observer. The wavefunction collapse is the only phenomenon that appears to provide a link between the neutral reality and the first and third person perspectives that derive from it. It is possible that, rather than the properties of reality themselves changing as a result of where it is observed/measured, as is implied by wave-particle duality, the *perception* of the neutral reality changes as a result of where it is observed/measured. The neutral reality is in its true form before it is observed/represented by a conscious observer. If von Neumann is correct, and the wavefunction collapse occurs in conscious observation, the superposed states could possibly *be* the neutral reality, and the collapse could potentially be the moment the neutral reality is translated into its represented forms, first and third person perspectives (Fig. 12.3). The superposition is consistent with the proposed neutral reality within RPNM. It can never be observed/accessed directly, as it collapses into a single state, or perspective representation, as soon as it is measured. However, it can and has been conceptualized, so there is no reason to believe we cannot proceed to Stage Three from this formulation. By understanding the precise nature of the wavefunction collapse (including how the wavefunction collapses, known as the measurement problem), we could potentially derive the elusive translative bridge laws. Assuming von Neumann, London, Bauer, and Wigner are correct, how and where is the wavefunction collapse facilitated in the brain, and additionally, what is a possible way quantum mind models can explain how the neutral reality is represented differently in the 'first person' perspective than in the 'third person' perspective?

Perhaps the most well known neurobiological model of consciousness that accommodates quantum effects is Roger Penrose and Stuart Hameroff's Orchestrated Objective Reduction (Orch-OR) hypothesis. Although von Neumann and Co. argued for consciousness collapsing the wavefunction, none of them explicitly proposed what consciousness itself was. Penrose's Objective Reduction model does just that. Penrose OR is a speculative interpretation of the wavefunction collapse that attempts to reconcile relativity with quantum theory (Penrose 1989). It hypothesizes that in addition to large objects, spacetime curvature exists among quantum superpositions as well, but in discrete bits. When the size exceeds the Planck scale (10^{-35}m), gravity begins to influence spacetime. Consequently the superposition collapses and spacetime becomes continuous. This is the objective reduction (OR) wavefunction collapse. The time taken for a superposition to collapse is described by Penrose's equation $E=h/t$, in which E is the energy of the superposed mass (directly related to its size), h is Planck's constant, and t is the quantum coherence time until objective reduction (the time a superposition takes to collapse). Rather than defining the wavefunction collapse as being caused by an external conscious observation, which is how von Neumann, London, Bauer, and Wigner described it, the Penrose OR hypothesis describes wavefunction collapse from another perspective: as a self-collapse due to gravity acting on spacetime. However, since Penrose agrees with von Neumann and Co.'s general interpretation – that conscious observation does cause collapse – he reasons that self-collapse must entail self-conscious observation – in other words, first person consciousness. The OR wavefunction self-collapse creates a discrete conscious moment of the single possibility/collapsed state. This provides the potential explanation for the differences in the 'first person' perspective and the 'third person' perspective representations of the neutral reality within RPNM (Fig. 12.4). Penrose OR collapse describes the process of self-observation or self-representation, when the neutral reality/superposition is represented/self-collapsed in the first person perspective. The von Neumann and Co. collapse describes the process of observation of external events (such as the results of the double slit experiment), when the neutral reality/superposition is represented/collapsed in the third person perspective. For example, a self-collapse within the brain would entail the first person representation of neutrality/superposition

A, what we call a mental event, and a collapse due to an external observation of the neutrality/superposition A causes the third person representation of neutrality A, what we call a physical brain event. Comprehending the nature of these wavefunction collapses could potentially allow us to find the underlying translative bridge laws between the neutral reality and our first and third person representations of it.

It is important to point out here that stating the von Neumann-London-Bauer-Wigner collapse entails the third person representation, and that the Penrose collapse entails the first person representation, does not imply are two different types of wavefunction collapse. Although Penrose himself never stated this explicitly, I believe the von Neumann and Penrose versions can be seen as simply two different ways of describing the same phenomenon, a wavefunction collapse in the observer's mind. As I have previously noted, the 'first' and 'third' person perspectives are also just linguistic terms and what can be seen as a 'first' or 'third' person perspective depends on one's frame of mind at any given moment. For example, I have emphasized numerous times that every representation can actually be seen as a first person representation with respect to the observer; an observer's 'third person' perspective of a particular neutrality in a subject's brain fact constitutes a first person perspective of a separate neutrality in the observer's own brain. Thus, no matter whether the representation is 'first' or 'third' person, it involves one type of wavefunction collapse within the observer's brain, which is facilitated by Penrose's Objective Reduction mechanism. Wigner deduced from his thought experiment that collapse is relative to the observer. Thus, a superposition/neutrality 'A' within a subject's brain when self-observed/self-collapsed (the subject plays the role of the observer, or rather self-observer, here) will result in the subject's first person representation of neutrality 'A', but when this subject's superposition/neutrality 'A' is observed as an external event in the 'third person', it will result in a wavefunction self-collapse of a separate superposition/neutrality 'B' within the external observer's brain. This will entail the 'third person' representation of 'A', which is identical to the first person representation of 'B'. The important point is that all these representations, whether 'first' or 'third' person, are the same type of wavefunction self-collapse within the observer's brain (Fig. 12.5).

So how are wavefunction collapses facilitated in the brain? The most experimentally consistent and well established brain-wide correlate of consciousness is gamma synchrony neural oscillation at 25-100Hz, typically at around 40Hz, which disappears during general anesthesia. Many neuroscientists originally abandoned pursuit of gamma synchrony as a global brain-wide neural correlate of consciousness, because the source of the synchronous oscillation could not be located within integrated axonal spike patterns. However recent studies have revealed that neuroscientists were previously looking in the wrong place for the source of the gamma synchrony; rather than axonal synapses – which do not give rise to neural coherence – it has been established that gap junctions between the dendrites of adjacent neurons are the mediators of gamma synchrony (Hameroff 2005). Webs of neurons interconnected via dendritic gap junctions form a unified and continuous internal cytoplasmic milieu bound by a common cell membrane, which fires synchronously "like one giant neuron". Stuart Hameroff hypothesizes that while axonal cross-communication arbitrates non-conscious neural computation, consciousness itself occurs in these dendritic webs (Hameroff and Penrose 1996). He claims that gamma synchrony originates from a coherent oscillation of dendritic membrane protein conformations. These protein conformations change due to the pi electron orientations in the proteins' hydrophobic pockets, which are selectively obstructed by anesthetic gases. However, this interpretation still leaves out the problem of the brain-wide synchrony oscillation with zero phase time lag, which is impossible in a classical mechanics model of dendritic communication. Hameroff concludes that quantum entanglement is responsible for the brain-wide gamma synchrony with instantaneous communication between all areas of the brain.

Hameroff proposes a biological framework involving computation within the microtubules that compose the neuronal cytoskeleton, specifically within dendrites (Fig. 12.6). The tubulin protein dimers (made of α and β monomers) which comprise microtubules contain hydrophobic pockets of aromatic hydrocarbon rings. These rings contain a large number of pi electron bonds, which resonate between one of two states, causing conformation change of the tubulin protein between two states. The coherent patterns of billions of tubulin proteins oscillating between two conformations act as a binary form of classical computation within microtubules. Hameroff however claims that the pi electrons within the hydrophobic pockets can also exist within a quantum superposition, and consequently, tubulin protein conformations exist in superposition between two states as well. When the superposition collapses via objective reduction into a single conformation, a discrete conscious moment/observation of the collapsed state occurs. These discrete moments of consciousness occur in rapid succession which gives the illusion that consciousness is a unified stream of uninterrupted thought. The gap junctions between dendrites of adjacent neurons creates a well regulated and unified internal system of communication through which all microtubules, connected by microtubule associated proteins (MAPs) to each other and membrane proteins like dendritic spines/synaptic receptors, within a neural network can cross-communicate to produce a coherent stream of conscious states, which is what gives rise to the gamma synchrony oscillation. Using Penrose's equation $E=h/t$, Hameroff calculated that the time it takes for a 3 nanogram superposed brain mass – an integrated microtubule network – to self-collapse is around 25 milliseconds. This time interval falls into the frequency range of gamma waves at 40Hz (40 oscillations per second). From this calculation Hameroff concludes that a discrete wavefunction self-collapse within the microtubules corresponds to a single oscillation in the gamma frequency range from 25-100Hz (for example, 1 second/0.025 seconds per self-collapse or oscillation = 40 self-collapses or oscillations per second). Since a wavefunction self-collapse is a conscious moment, this means that each oscillation of the gamma wave corresponds to a discrete conscious moment. Hameroff mentions that ancient Buddhist texts have documented highly experienced meditators who claimed to be able to count the number of discrete conscious moments they experienced while in a deep meditative state (Hameroff 2011). These meditators counted approximately 6.5 million conscious moments within a 24 hour period. This is equivalent to around 75 conscious moments per second, or 75Hz. This falls within the higher frequency range of gamma wave oscillations, which the Lutz et al. experiment demonstrated were markedly stronger in highly trained monks during their 'objective compassion' meditative state (Lutz et al. 2004).

A number of problems have arisen with Hameroff's specific microtubule model. First, neuronal microtubules consist of 'B' lattices rather than the 'A' lattices more apt for computation (Kikkawa et al. 1994). Additionally Max Tegmark calculated that at brain temperatures quantum coherence states could only last on a femtosecond scale, far less than the 25 milliseconds required by Orch-OR to facilitate the 40Hz synchrony (Tegmark 2000). Notwithstanding the concerns with Hameroff's hypothesis, a quantum basis for mind, including Penrose's Objective Reduction, remains viable. Contrary to prior belief that biological systems were too warm for quantum effects, quantum coherence has been found to play a role in a number of biological mechanisms, including electron energy transfer in photosynthesis (Engel et al. 2007). Others have suggested alternate neurobiological mediators for the quantum mind. Henry Stapp proposes quantum collapses within the synapses (Quantum Approaches to Consciousness 2013), and Danko Georgiev suggests that actin filaments, not microtubules, are the mediators for quantum computation (Georgiev 2006). One distinction that sets the Quantum Mind hypotheses apart from any other model of the mind is that thus far, they are the only models that provide a possible a way for us to proceed to Stage Three, establishing true translative bridge laws between the neutral reality and our first and third person representations of it, rather than making identity claims, drawing correlations, or modeling consciousness as a general process or system.

13. Possible Biological Experiment to Test for the 'Conscious Observation Causes Wavefunction Collapse' Hypothesis

Fred Thaheld has proposed a possible experiment to test von Neumann, Penrose, et al.'s proposition that conscious observation causes wavefunction collapse (Thaheld 2005). Thaheld proposes testing the wavefunction collapse capabilities of the photoreceptors in the eyespot of *Euglena gracilis.* Since *E. gracilis* is unicellular, the assumption is that it is not conscious, at least not in the same way as a human brain. When the eyespot pigment molecules detect photons of ~500nm wavelength, they convert it to electrical signals which are transmitted to the flagellum. This causes a flagellar movement known as photomotion. The experiment is set up such that a superposed photon state is detected by an *E. gracilis* eyespot before it is detected by the human observer in the experiment. If photomotion occurs when the superposed state is detected by the eyespot, it would indicate that the eyespot had collapsed the wavefunction into a single state photon wavelength, causing the electrical signaling and flagellar motion. This would imply that conscious observation is not necessary for wavefunction collapse, and that non-conscious biological organisms can observe/measure and collapse the wavefunction as well. However, there are several problems with this experiment. First, according to Orch-OR, *E.*

gracilis may not have the complex consciousness of humans, but self-collapsing wavefunctions within the microtubules of the cell would be enough to cause a primitive consciousness. So it could be argued that E. gracilis is in fact a conscious observer. Second, even if E. gracilis is not conscious, it would still be analogous to the measuring device in the original double slit experiment. Let's apply the Wigner's Friend thought experiment and the 'von Neumann chain' to this scenario. Where does the wavefunction collapse really occur? In the eyespot when the superposed photons are detected or in the *E. gracilis* flagellum when the human observes to see if it is engaging in photomotion? In both of these cases we are inserting the measurement within the external world and splitting the universe into superposed and collapsed states, which as von Neumann indicated was illogical. From this we can see that if there is no conscious observer to see the results of the experiment the photons and the flagellum remain in a superposed state. The only location the collapse can logically take place is still where the final measurement of the experiment is occurring, the conscious brain of the human observer. So it remains to be seen whether a well set up and rigorous experiment can really test the 'consciousness causes collapse' hypothesis.

14. Evidence from Fringe Domains That Support the Quantum Mind Hypotheses

In addition to the above evidence, collective evidence within certain fringe domains within the past decades has also suggested a quantum basis for mind. This evidence is far from conclusive, but is much more prevalent than most skeptics will admit. I will not go into this evidence in detail but I feel it does need to at least be briefly mentioned as it has been unjustly dismissed out of hand by the majority of scientists who are a priori convinced of its inaccuracy.

Perhaps the most famous experiment within the field of parapsychology is the Ganzfeld test (Palmer 2003). A subject is given complete sensory deprivation: covered eyes and earphones with white noise. His partner sits in another room, and is given a panel of four possible images to 'mentally project' to the subject. The subject tries to describe the images he sees, and after the experiment is shown the four images. He then picks the image that best fits what he saw. By chance we would expect exactly 25% success rate, but after thousands of independently replicated studies, the success rate hovers at around 27%. This difference seems small, but is statistically significant.

Remote viewing was funded by the US government from 1970 to 1995, after which the technique was largely discredited. However, experimental neuroscientist Michael Persinger obtained interesting results when testing former US employed remote viewer Ingo Swann (Persinger et al. 2002). Swann was asked to describe, using drawings, certain obstructed and distant images not revealed to him until afterwards. His drawings were evaluated for their accuracy based on a scale correlating specific aspects of his drawings with the actual obstructed image. Swann was determined to be around 40% accurate, incredibility statistically significant. A control group of 40 individuals was also tested, and did not score nearly as well in accuracy. Additionally, fMRI scans showed increased activity in Swann's occipital, temporal and frontal lobes, which did not occur in the control group. Psychologist and Committee for Skeptical Inquiry fellow Richard Wiseman even stated upon reviewing the evidence, "I agree that *by the standards of any other area of science that remote viewing is proven*, but begs the question: do we need higher standards of evidence when we study the paranormal? I think we do" (Penman 2008).

Skin conductance tests have revealed that a subject exposed to random images flashed on a computer screen, mostly benign and occasionally upsetting, would anticipate the content of the image about to appear (skin conductance increases as a result of a mental shock stimuli), without possibly having been able to know about the content beforehand (Spottiswoode and May 2003). The results have been repeated on subjects tested with audio stimuli with occasional startle noises.

One of the more fascinating studies within parapsychology, which is seldom mentioned even within that discipline, concerns the case studies of young children who claim to remember lives of previously deceased individuals (Stevenson 1993). Child psychologist Ian Stevenson was faced with many such cases of children claiming to remember past lives. Becoming intrigued, he tried to verify these statements, and eventually traveled around the world, documenting over 2500 cases of this type. In the approximately 250 cases where confirmation was possible these statements did in fact correspond to actual previous lives of ordinary people (not well known individuals of whom much information about their lives could be readily gathered), in many cases which the child's family had not been previously acquainted with. Perhaps most intriguingly, Stevenson documented more than 100 cases in which unusual birthmarks and birth defects corresponded to unusual, and often fatal, wounds on the 'previous life'. These birthmarks and birth defects corresponded with a high degree of accuracy autopsy reports of the individual whose life the child claimed to remember.

What does this evidence imply? If true, this would certainly indicate that mind is non-local, not only within individual brains but between separate brains as well, resulting in a transfer of mental information between two brains at an instantaneous, or zero phase time lag scale (I refrain from using 'psychic' phenomena, as that is a loaded term). The only known example of this type of non-local interaction occurs in quantum entanglement, which has been demonstrated by Bell's Theorem (Felder, 1999). However, no known interpretation of quantum entanglement allows for the type of information transfer which would be necessary for these non-local mental phenomena to be true. Nevertheless, it is unwise to dismiss this evidence out of hand simply because we lack a scientific framework with which to explain it. If true, these phenomena will not prove the existence of magic; a naturalistic explanation will be forthcoming, and may take us a step closer into understanding the nature of the mental, the physical, and their relationship to the neutral reality.

15. Final Thoughts

RPNM gives a great number of postulations, many of them untestable through our current scientific means. Although there is no grand single experiment that could falsify RPNM, as the pieces start to come together in the different scientific domains, we will begin to see whether this paradigm has any validity. The problem of consciousness is an enormous one. We often lose sight of the big picture by narrowing focus on a very small area of inquiry in a specific scientific domain. A significant shortcoming in the study of consciousness is the lack of cross-communication among the diverse groups of domains investigating this phenomenon. It is imperative to draw evidence from an extensive range of scientific disciplines, including research investigating fringe hypotheses. Historically, solutions to the most pressing scientific problems (e.g. natural selection) were uncovered not by narrowing focus on a specialized area, but by integrating ideas from all fields of study.

Figures

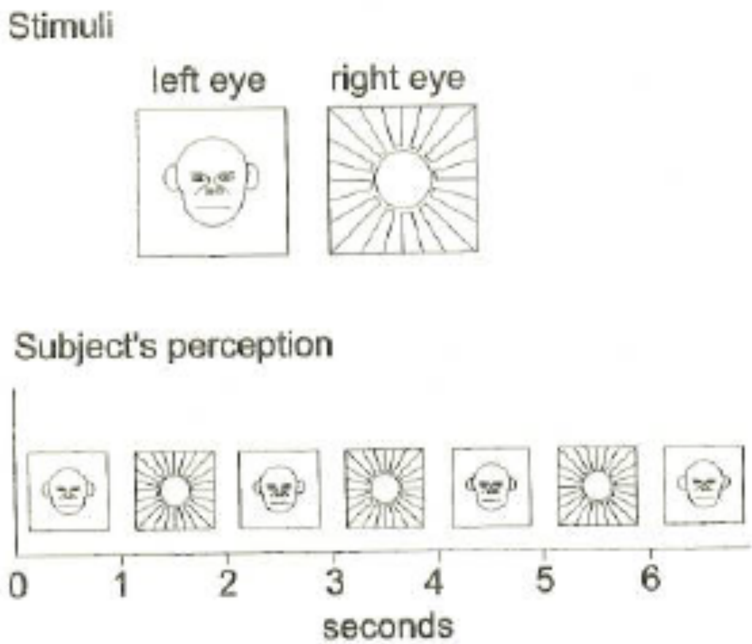

Figure 1.1. Binocular rivalry, used to test neural correlates of visual awareness. The subject is presented with two different images in the left and right eye. Instead of experiencing both images at the same time, the subject has a stable, alternating visual awareness of one or the other. (From Churchland 2002.)

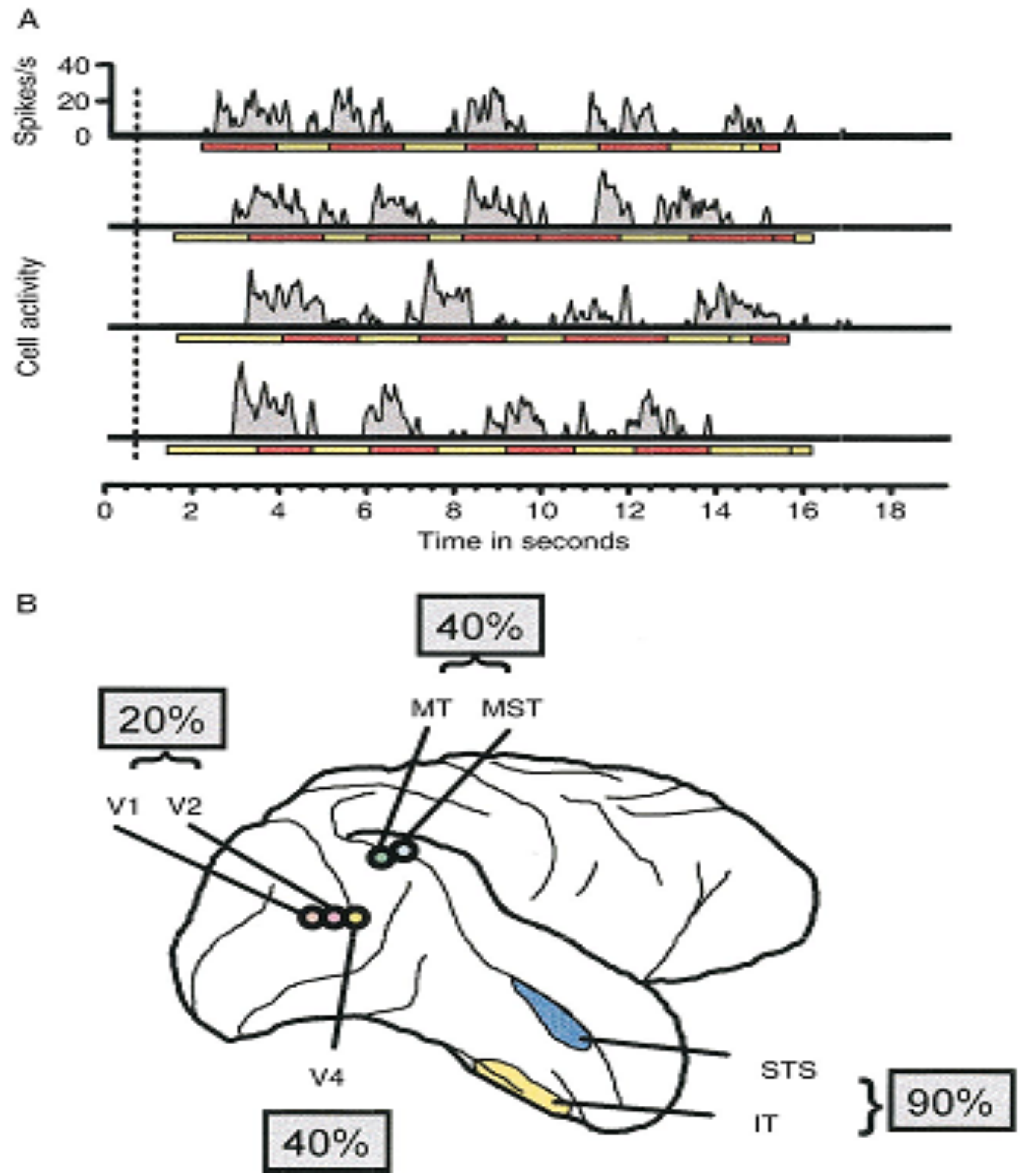

Figure 1.2. Logothetis binocular rivalry experiment. (a) Microelectrode measurement of activity of single neuron in V4. The red bars represent when the monkey reported visual awareness of the image in the right eye (the face image); the yellow bars represent visual awareness of the left eye image (the non-face image). This particular neuron becomes active only when the monkey is aware of the face image. (b) Activity of visual cortex neurons when change in visual awareness is reported. 90% of ITC and STS cells, 40% of V5/V4 cells, and less than 20% of V1/V2 cells became active only when *awareness* of the face image was reported; the other percentage remained active as long as the face visual stimulus was present. (From Leopold and Logothetis 1999.)

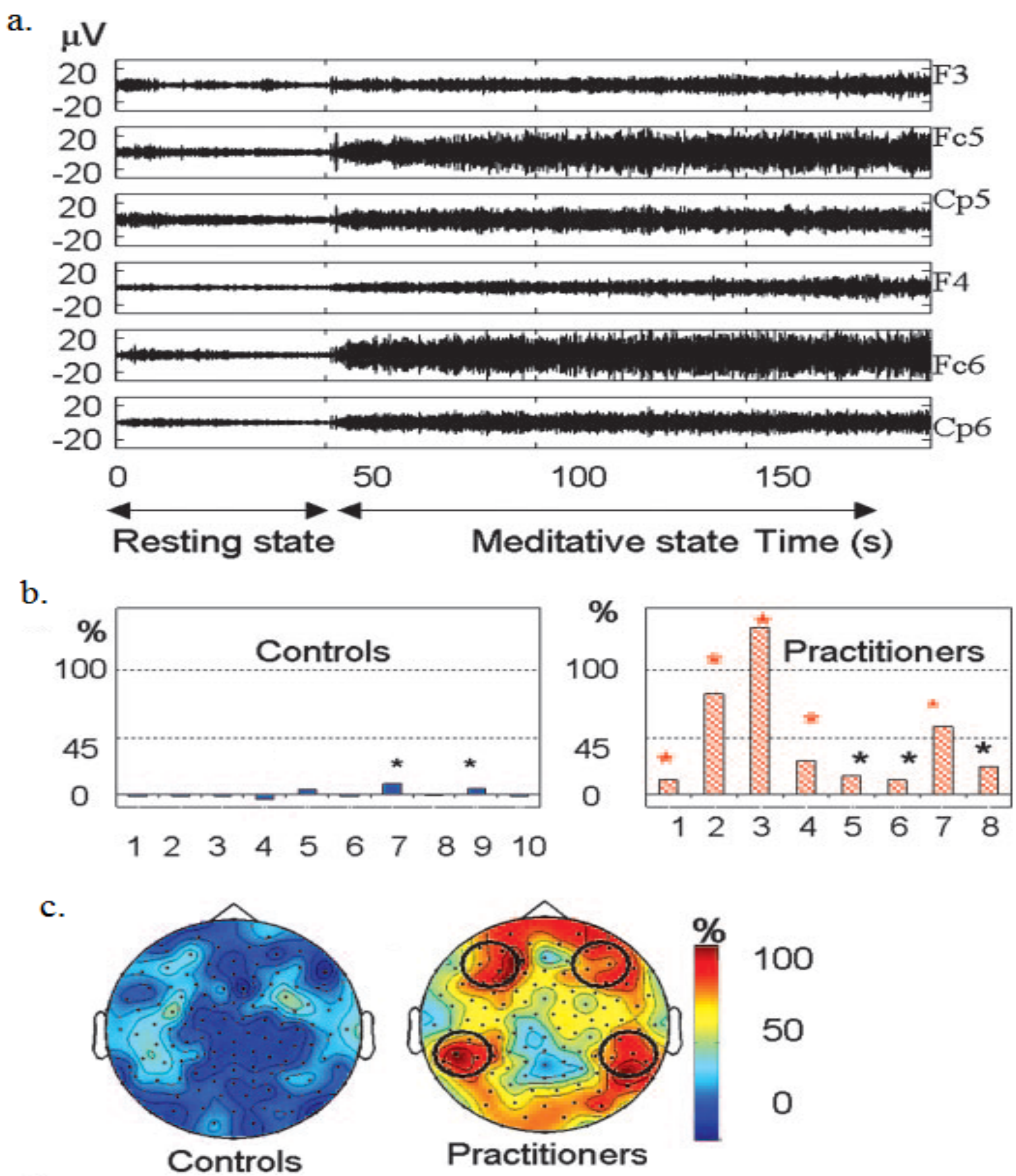

Figure 1.3. Meditation and gamma synchrony. (a) Raw EEG of Buddhist practitioner. The 'resting state' represents the normal meditative state. At t=45s the subject was instructed to generate feelings of "objective compassion", beginning the 'meditative state'. (b) The abscissa represents the subject numbers; the ordinate represents the difference in gamma wave intensity between resting state and "objective compassion" meditative state. One can clearly see that the Buddhist practitioners' gamma power increases tremendously from resting to meditative state, whereas the controls' gamma power exhibits little change. (c) Distribution of gamma activity throughout the scalp, as measured by EEG. The practitioners exhibit much stronger brain-wide, long range global gamma synchrony, from the frontal to parietal and temporal lobes. (From Lutz et al. 2004.)

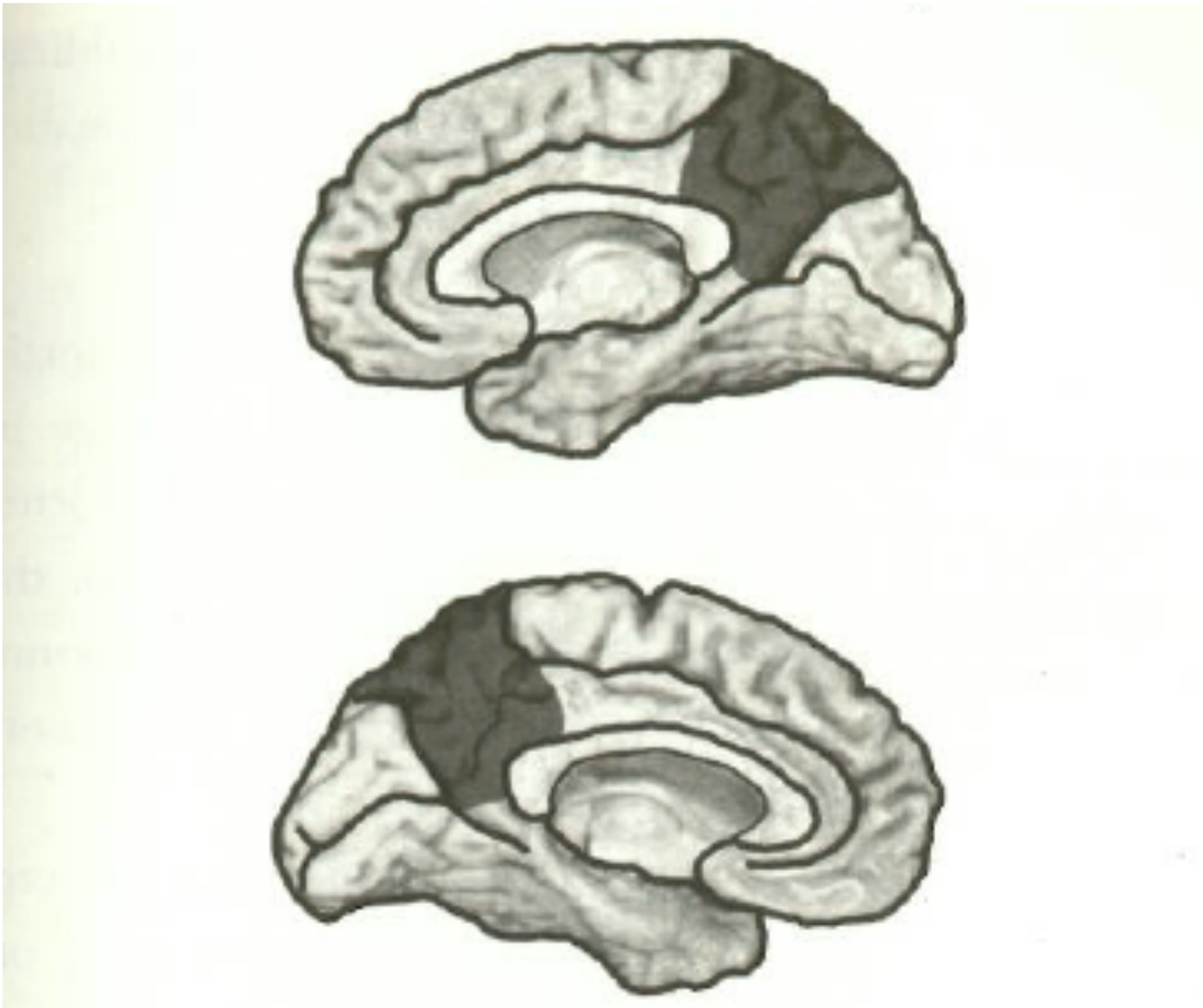

Figure 1.4. Location of the posteromedial cortices (PMCs) in the brain. The PMCs (indicated by the shaded areas) are essential in Damasio's model of constructing an autobiographical self. (From Damasio 2010.)

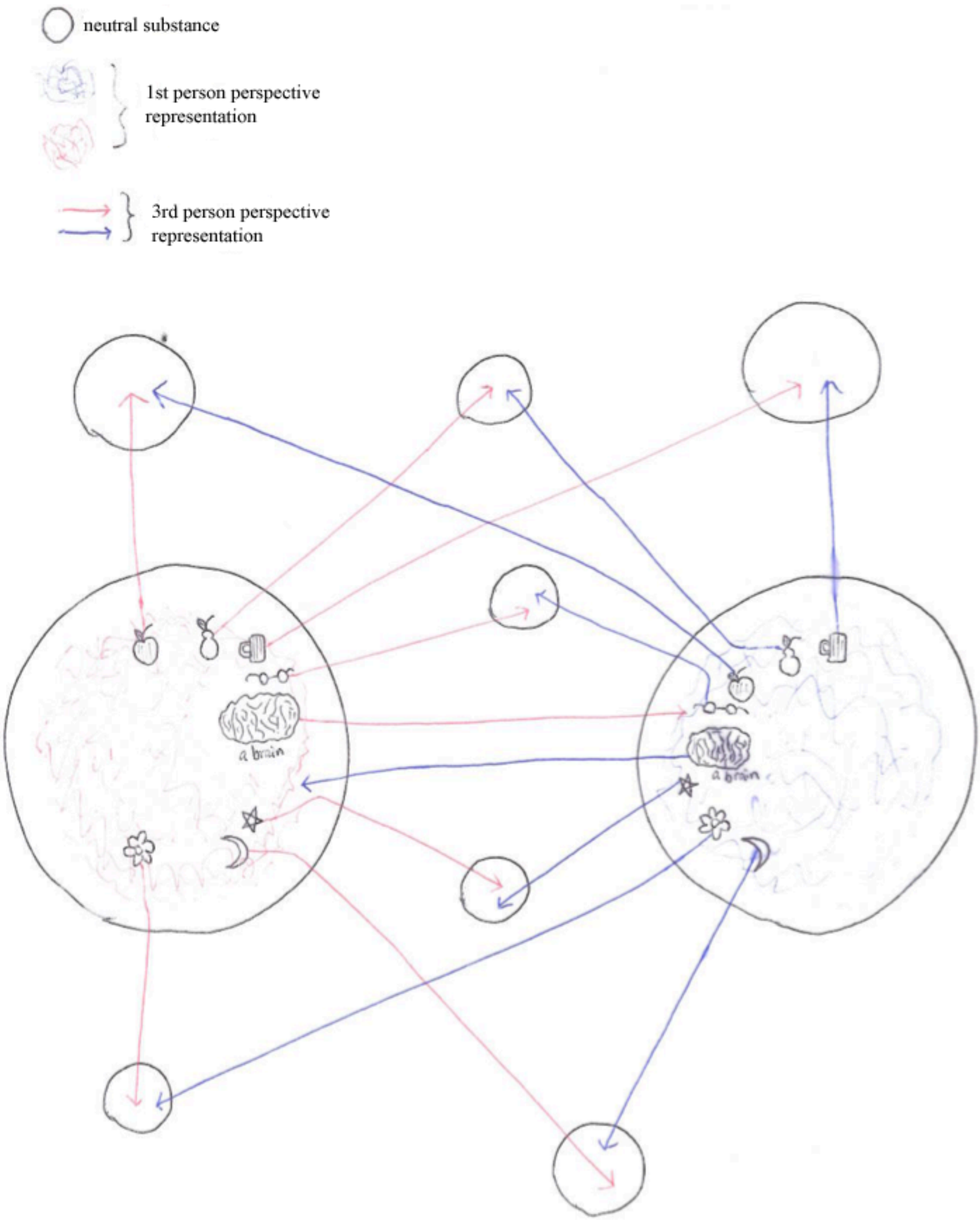

Figure 3.1. Relational Perspective Neutral Monism (RPNM). Reality is a neutral substance/neutrality (represented by circles). 'Brains' are neutralities as well (represented by the two large circles). They perceive the external world/other neutralities in the third person (represented by arrows) and themselves in the first person. Note how they perceive each other in the third person perspective, as physical brains.

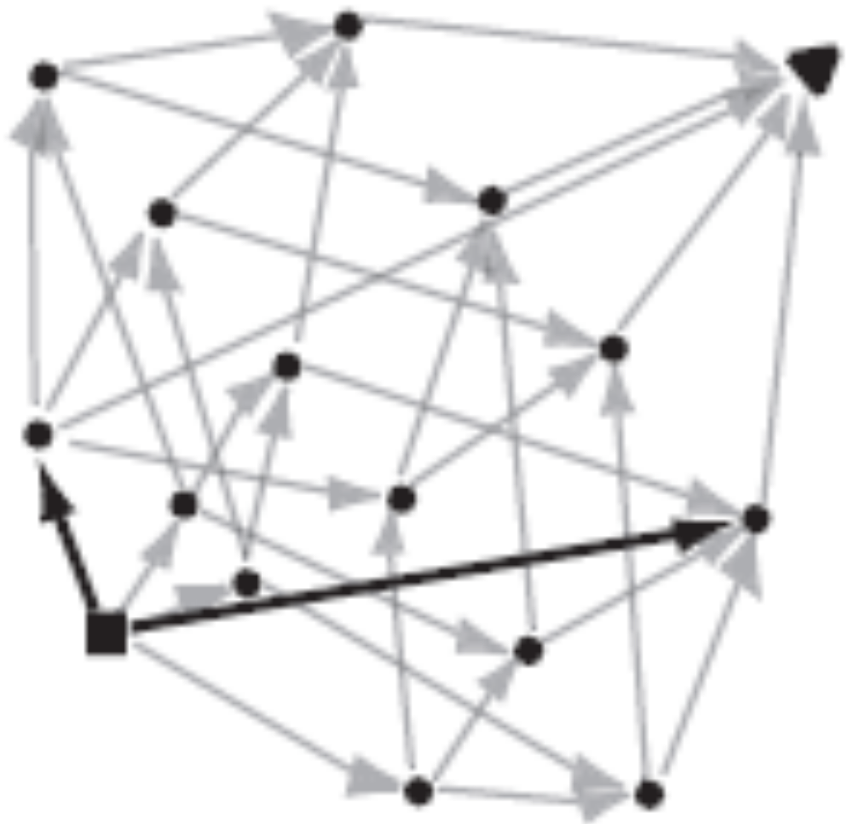

Figure 5.1. Tononi's multidimensional qualia shape, or "quale", used to model a particular subjective experience. The shape has been flattened into a two dimension approximation. The arrows represent various neural cross-communications/firings, the points on the shape indicate specific qualitative aspects of a subjective experience. Quale similar in shape are expected to be similar qualitative experiences. (From Tononi 2008.)

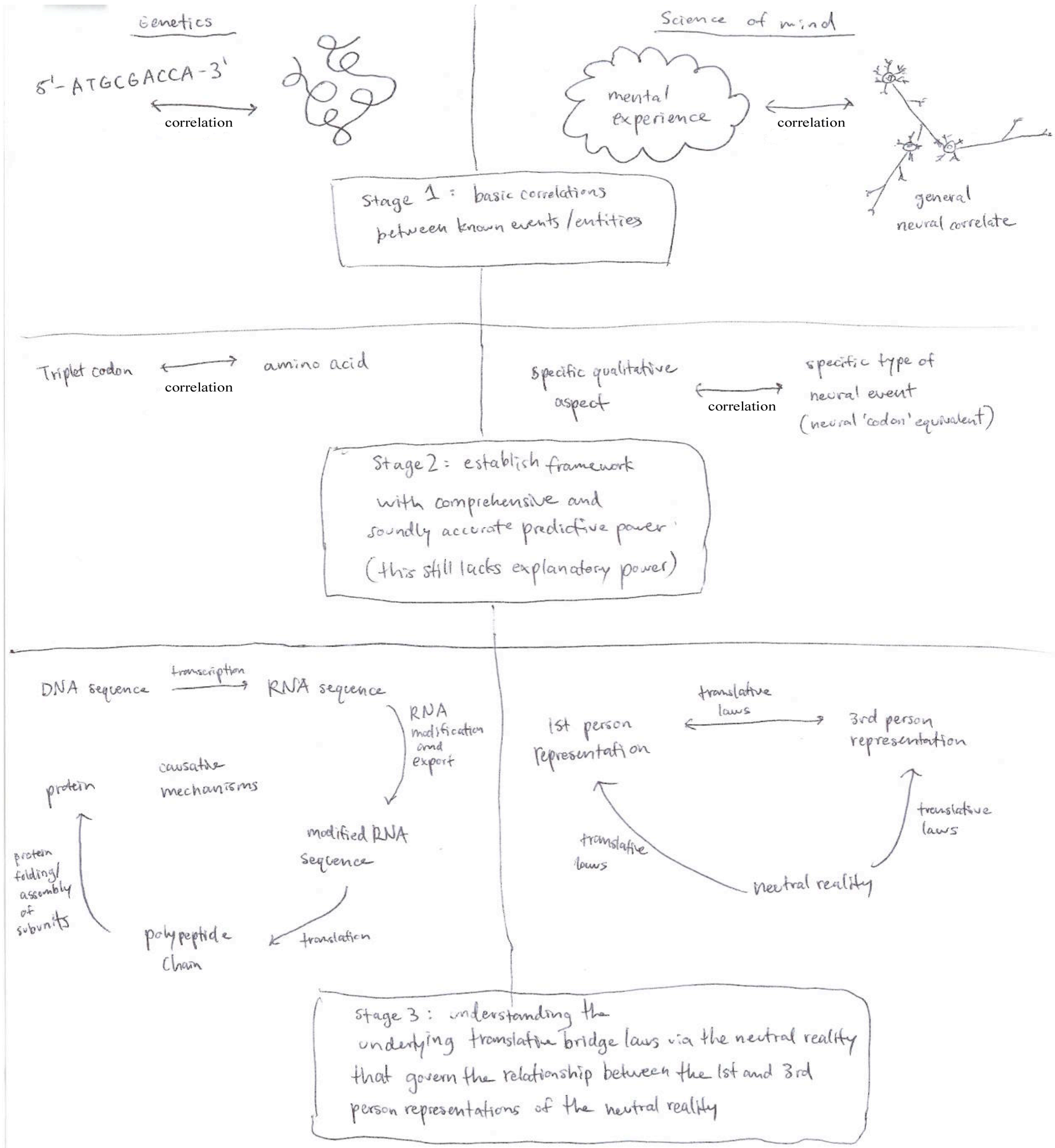

Figure 7.1. Three stages to developing a science of mind (with an analogy to the science of genetics). 1. Establish basic correlations. 2. Construct algorithm to translate specific types of neural events to specific qualitative aspects, and thus be able to predict novel experiences. 3. Understand the true underlying translative bridge laws via the neutral reality that that govern the relationship between the first and third person representations of the neutral reality. Note that the previous two steps are like scratching the surface compared to Stage Three – drawing correlations between perspectives but not understanding the underlying neutral reality from which both perspectives are derived. Only Stage Three, by understanding the neutral reality, can get us beyond correlations.

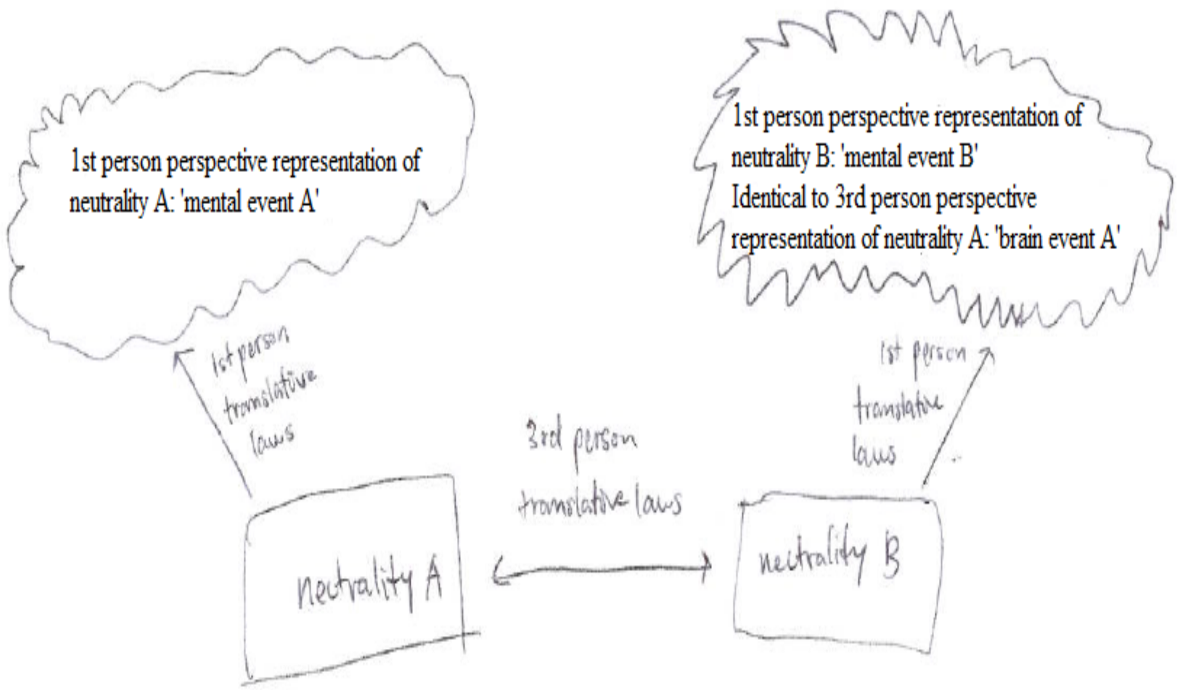

Figure 7.2. The actual translative relationship between the 'mental', 'physical', and neutral. Neutrality A translates to a mental event (via the first person translative laws) when represented in the first person. When represented in the third person however, first it translates (via the third person translative laws) into a separate neutrality B, which then translates (via first person translative laws) into a brain event. This occurs because my third person representation of *your* neutrality as a brain event is also *my* mental event. So in order to translate between the two perspectives, first we would have to translate a mental event A into its corresponding neutrality A (which is activated when A is first person representing itself), then translate neutrality A into corresponding neutrality B (activated when B is third person representing A), and finally translate neutrality B into its corresponding mental event B (activated when B is first person representing itself), which is represented in the third person as 'the physical brain event corresponding to mental event A'. In other words, the first person representation of B (mental event B) is identical to the third person representation of A (brain event A), and thus the first and third person representations of neutrality A can also be stated as 'mental event A' and 'mental event B', respectively.

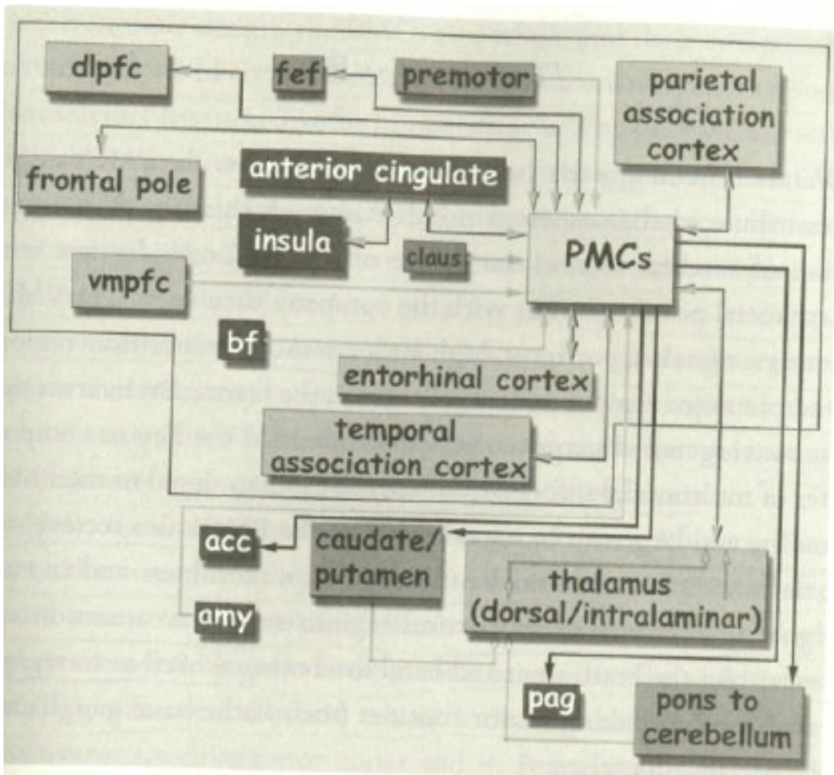

Figure 8.1. Diagram sketching neural cross-communications (converging and output signals) between the PMCs and other areas of the brain. (From Damasio 2010.)

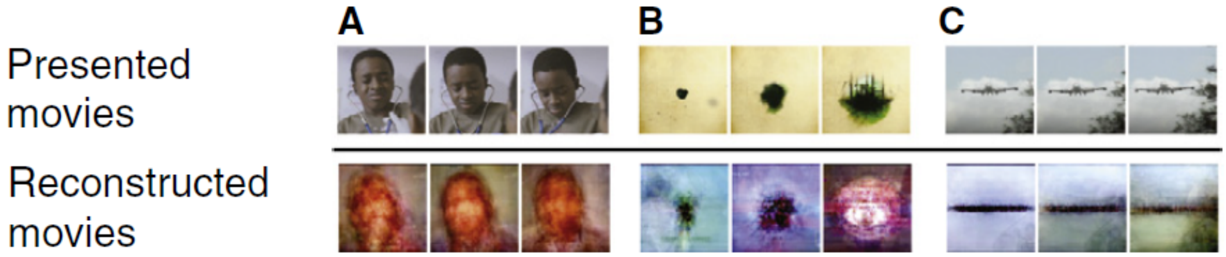

Figure 10.1. Reconstruction of visual experiences from brain activity. Using an algorithm that correlates specific qualitative aspects, e.g. shapes, colors, movement, with specific types of brain activity, a computer was able to roughly reconstruct any visual image/movie a subject was experiencing, by monitoring his brain activity using fMRI while he was viewing the images/movies. A, B, and C are three examples of this image reconstruction. The top row shows screenshots of the actual movies the subject was watching, and the bottom row shows screenshots of the computer reconstructed movies based on the subject's brain activity. (From Nishimoto et al. 2011.)

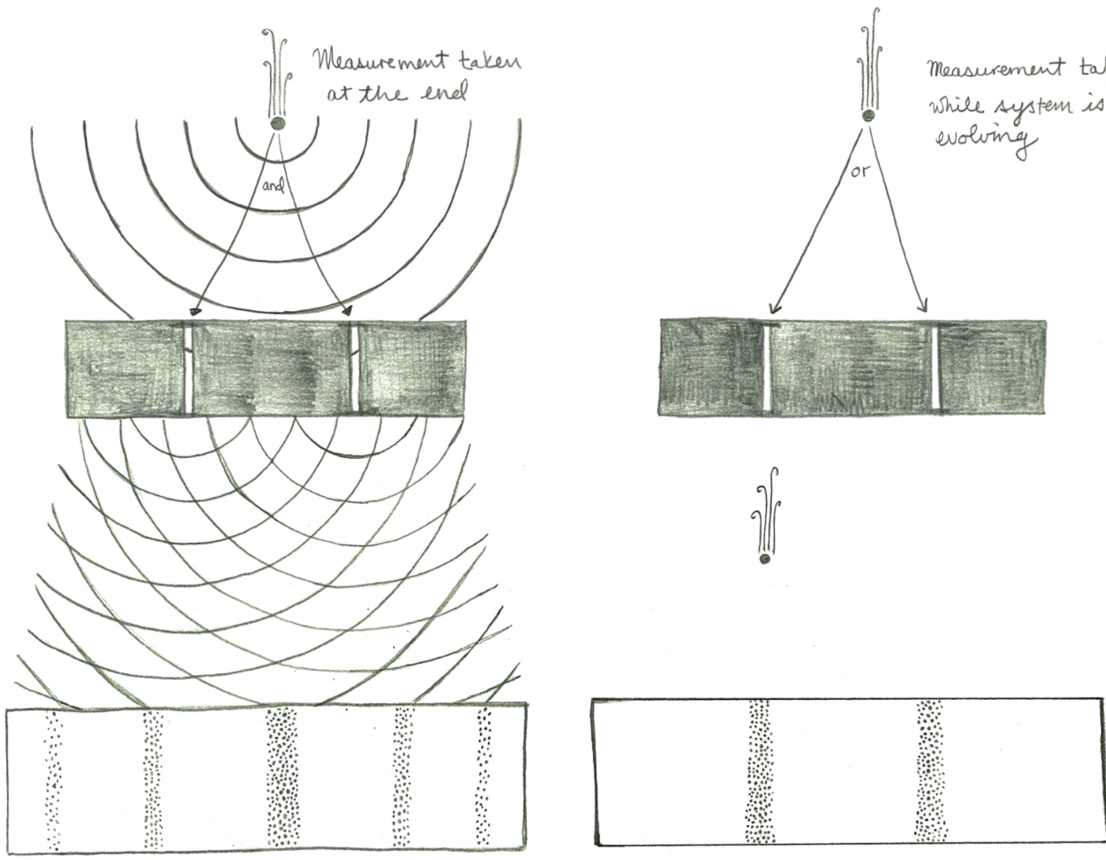

Figure 12.1. The double slit experiment that demonstrated wave-particle duality. Electrons are fired through the double slits at a photographic receptor plate. If no measurement is taken after the electrons enter the double slits, a wavelike interference pattern develops on the plate. If measurement is taken before the electrons enter the double slits, a two slit particle-like pattern develops on the plate. This implies that as soon as measurement is taken (measurement is taken on the photographic plate in the first experiment, and before the electrons enter the double slit in the second experiment), reality goes from behaving like a wave to behaving like a particle. (From http://www.oist.jp/photo/double-slit-experiment.)

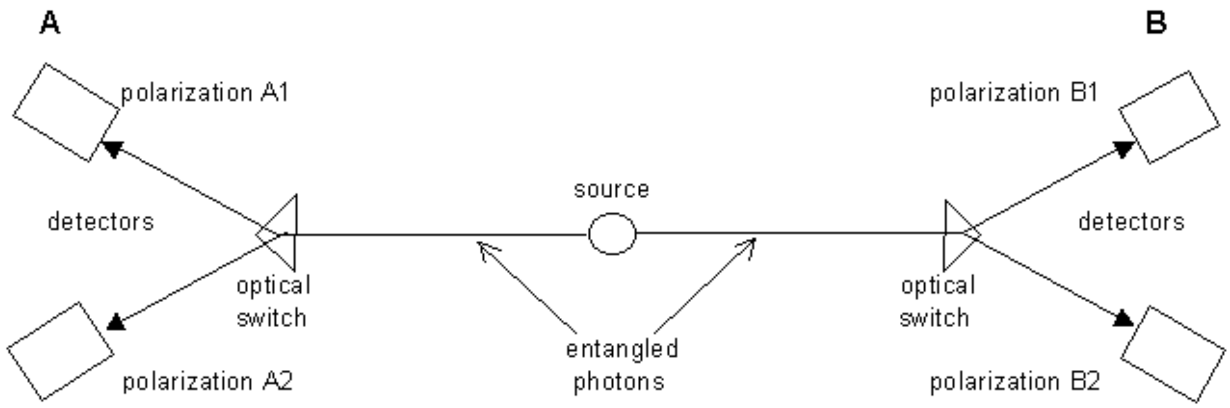

Figure 12.2. Bell's Theorem. A particle decays into two quantum entangled particles A and B with opposite charges/polarizations. They are both in a superposition of either charge until particle A is measured with an optical switch. That particle then collapses into a single charge/state, which is identified by detectors. However, as soon as particle A collapses, particle B, which has been separated from particle A and therefore should still be in a superposition, collapses into the opposite charge of A. This implies a non-local, instantaneous, faster than light-speed transfer of information. (From http://faculty.virginia.edu/consciousness/new_page_7.htm.)

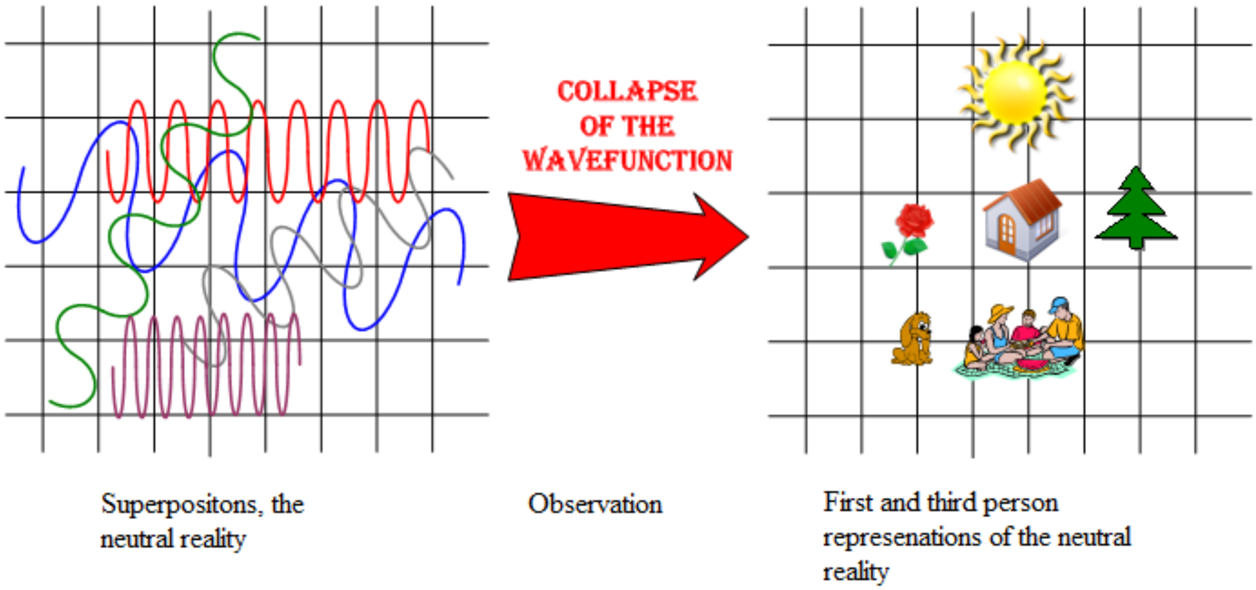

Figure 12.3. Wavefunction collapse provides possible translative laws from the neutral reality to the first and third person perspective representations. The neutral reality is the superposition, which collapses into a discrete state, or representation, when observed. (Based on http://bulldozer00.com/category/quantum-physics-2/.)

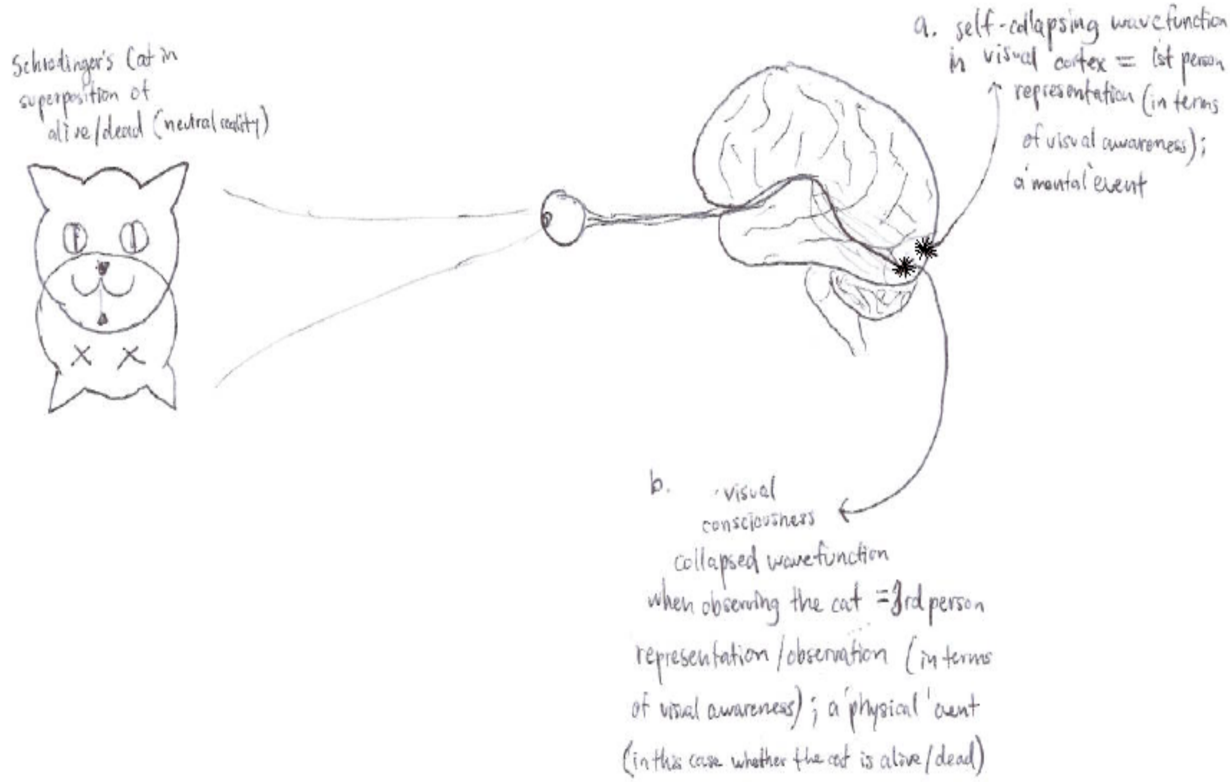

Figure 12.4. The difference between first and third person perspective representations, using visual awareness as an example. (a) A superposition/neutrality within the visual cortex self-collapses, which results in self-representation or the first person representation, a 'mental' event. (b) When a superposition/neutrality external to the observer is observed, the wavefunction is collapsed by the visual conscious observation, which results in the third person representation, or a 'physical' event.

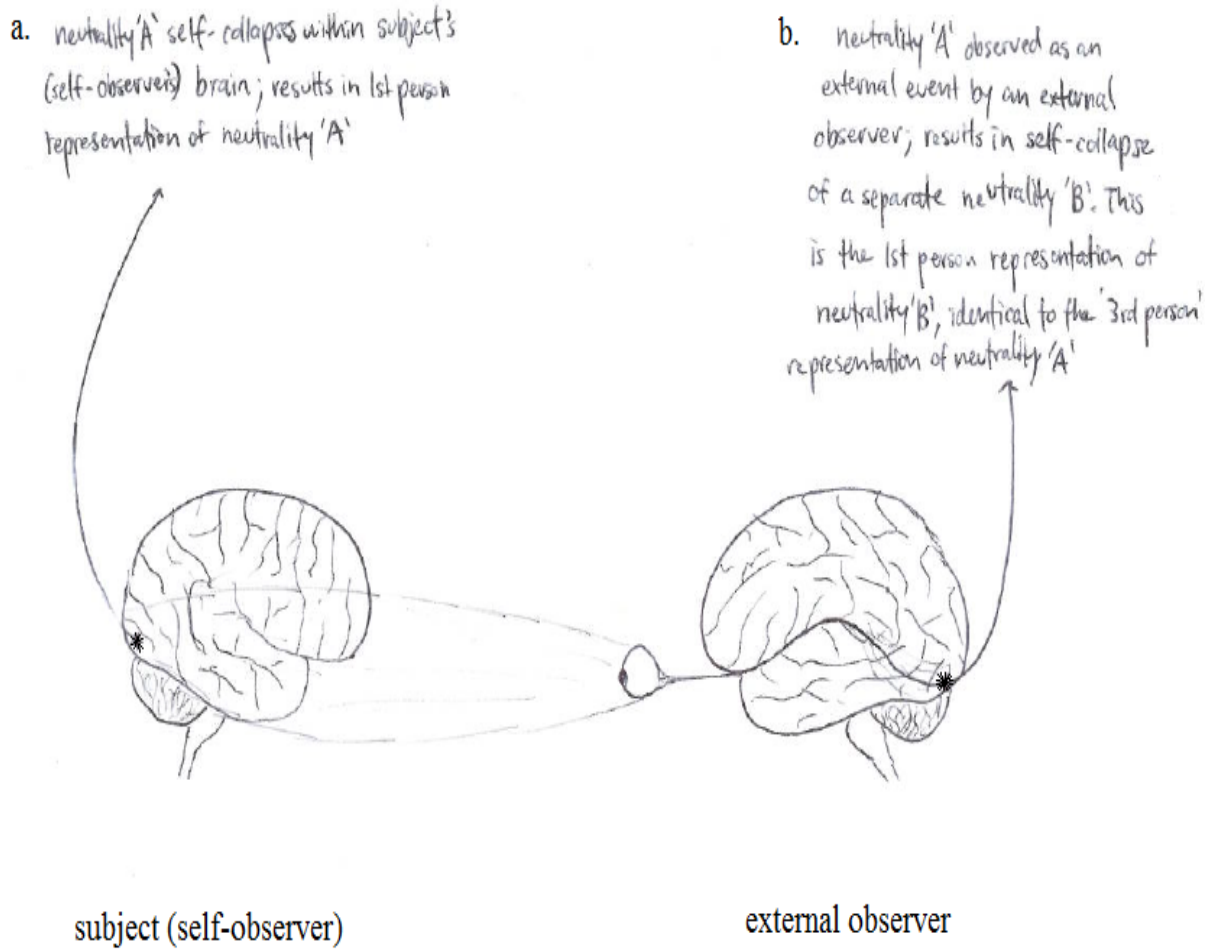

Figure 12.5. The von Neumann-London-Bauer-Wigner and Penrose versions of the wavefunction collapse are just different ways of describing the same phenomenon, a wavefunction collapse in the observer's brain. Wavefunction collapses relative to the observer, so it always occurs in a conscious observer's brain. (a) A superposition/neutrality 'A' self-collapses in the subject's (considered the observer or self-observer here) brain – via first person translative bridge laws – resulting in the first person representation of 'A'. This is how Penrose described the collapse. (b) When neutrality 'A' is observed by an external observer, the external observer's corresponding separate superposition/neutrality 'B' – via third person translative bridge laws – self-collapses in the external observer's brain – via first person translative bridge laws. This results in the first person representation of 'B', which is identical to the 'third person' representation of 'A'. This is how von Neumann, London, Bauer, and Wigner described the collapse.

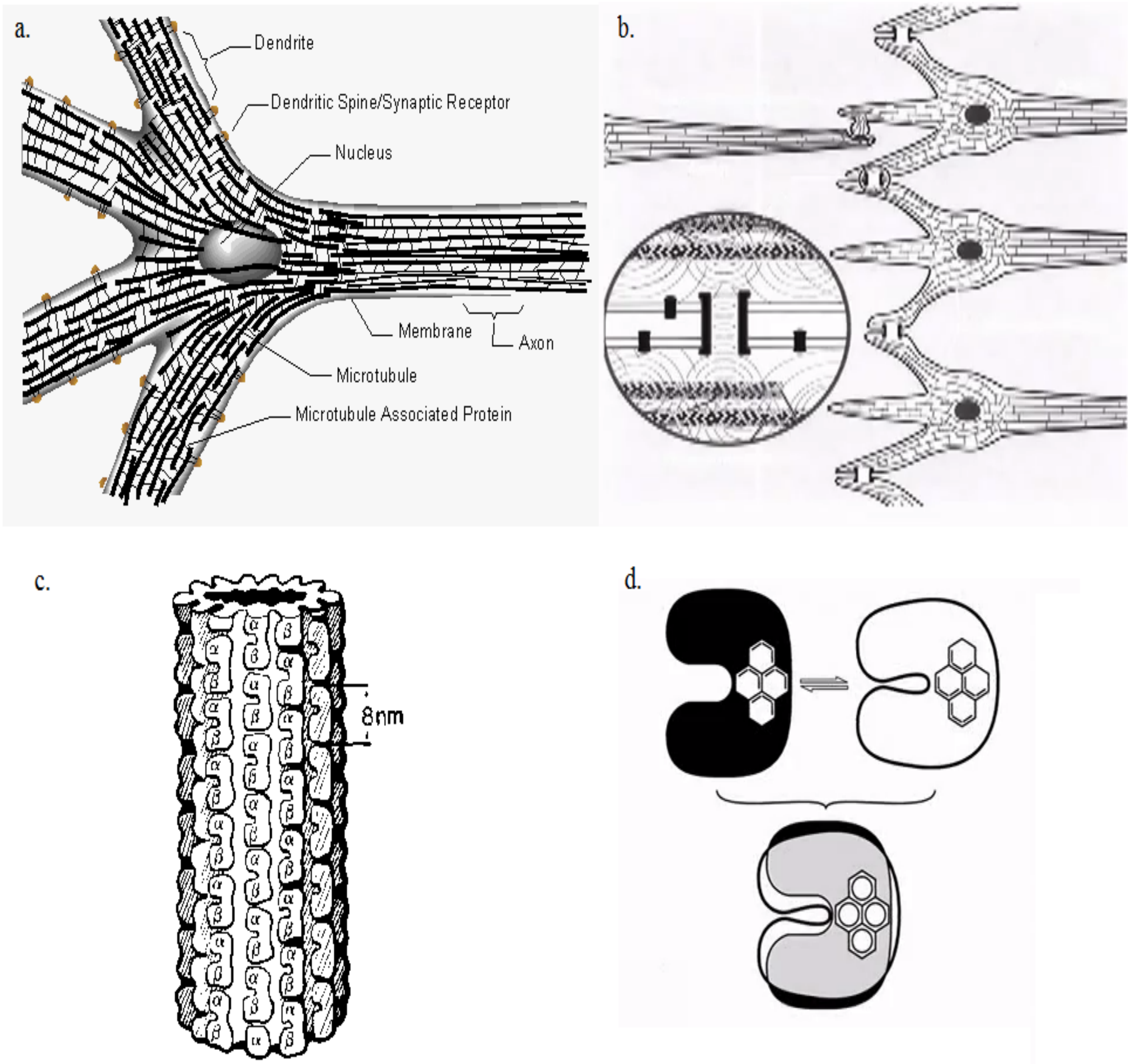

Figure 12.6. Hameroff's microtubule model for quantum computation. (a) Neuronal microtubules are highly interconnected by MAPs to each other and to membrane proteins, which can facilitate computation. (b) Webs of adjacent neurons connect via dendritic gap junctions to form an extensive isolated internal cytoplasmic milieu surrounded by a single membrane, capable of facilitating broad scale computation (c) Microtubules are composed of 8nm tubulin protein dimers, made of α and β monomers. Billions of interacting and cross-communicating tubulin proteins oscillating between one of two conformations acts as a form of binary classical computation within microtubules. (d) The pi bond electrons in hydrophobic pockets resonate, causing one of two protein conformations. Hameroff proposes that the pi bonds also exist in a superposition of the two states, causing the tubulin protein to be superposed as well. When the superposition reaches threshold it collapses, evoking a conscious moment of the collapsed state. Billions of interacting tubulin proteins collapsing coherently into one of two conformations creates the stream of conscious states within neuronal microtubules. (From Hameroff, Penrose 1996.)

References:

"Phenomenology." *(Stanford Encyclopedia of Philosophy)*. http://plato.stanford.edu/entries/phenomenology/, April 2013.

"Quantum Approaches to Consciousness." *(Stanford Encyclopedia of Philosophy)*. http://plato.stanford.edu/entries/qt-consciousness/, April 2013.

"The Mind/Brain Identity Theory." *(Stanford Encyclopedia of Philosophy)*. http://plato.stanford.edu/entries/mind-identity/, April 2013.

Alkire, M.T. and Miller, J. (2005). "General anesthesia and the neural correlates of Consciousness." *Progress in Brain Research* 150, 229-244.

Anscombe, Elizabeth. (2001). *An Introduction to Wittgenstein's Tractatus (Wittgenstein Studies).* South Bend, IN: St. Augustine's Press.

Bértolo, H., Paiva, T., Pessoa, L., Mestre, T., Marques, R., and Santos, R. (2003). "Visual dream content, graphical representation and EEG alpha activity in congenitally blind subjects." *Cognitive Brain Research* 15, 277-284.

Bohm, David. (1990). "Beyond Limits - A Full Conversation With David Bohm." Interview.

Bohr, Niels. (1958). "Discussion with Einstein on Epistemological Problems in Atomic Physics." *Uspekhi fizicheskikh nauk* 66, no. 4, 571-598.

Chalmers, David J. "Phenomenal Concepts and the Knowledge Argument." http://consc.net/papers/knowledge.html, April 2013.

Churchland, Patricia S. (2002). *Brain-wise: Studies in Neurophilosophy*. Cambridge, MA: MIT Press.

Crick, Francis. (1994). *The Astonishing* Hypothesis: *The Scientific Search for the Soul.* New York: Charles Scribner's Sons.

Damasio, Antonio. (2010). *Self Comes to Mind: Constructing the Conscious Brain.* New York: Random House, Inc.

Dennett, Daniel. (1988). "Quining Qualia." In: Marcel, A. and Bisiach, E. (eds). *Consciousness in Contemporary Science*, 42-77.

Engel, G.S., Calhoun, T.R., Read, E.L., Ahn, T.K., Mancal, T., Cheng, Y.C., Blankenship, R.E., and Fleming, G.R. (2007). "Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems." *Nature* 446, no. 7137, 782–786.

d'Espagnat, Bernard. (2005). "Consciousness and the Wigner's Friend Problem." *Foundations of Physics* 35, no. 12, 1943-1970.

Felder, Gary. (1999). "Spooky Action at a Distance." *Bell's Theorem: Physics*. http://www4.ncsu.edu/unity/lockers/users/f/felder/public/kenny/papers/bell.html, April 2013.

Georgiev, Danko. (2006). "Falsifications of Hameroff-Penrose Orch OR Model of Consciousness and Novel Avenues for Development of Quantum Mind Theory." *NeuroQuantology* 5, no. 1, 145–174.

Gold, Ian. (1999). "Does 40-Hz Oscillation Play a Role in Visual Consciousness?" *Consciousness and Cognition* 8, 186–195.

Hameroff, Stuart. (2005). "Breakthrough study on EEG of meditation." *Science and Consciousness Review*.

Hameroff, Stuart. (2011). "Dr. Stuart Hameroff On Quantum Consciousness and Moving Singularity Goal Posts." Interview.

Hameroff, Stuart and Penrose, Roger. (1996). "Orchestrated Objective Reduction of Quantum Coherence in Brain Microtubules: The "Orch OR" Model for Consciousness." *Mathematics and Computers in Simulation* 40, 453-480.

Hobson, J. A., Pace-Schott, E. F., and Stickgold, R. (2000). "Dreaming and the brain: toward a cognitive neuroscience of conscious states." *Behavioral and Brain Sciences* 23, 793–1121.

Horikawa, T., Tamaki, M., Miyawaki, Y., and Kamitani, Y. (2013). "Neural Decoding of Visual Imagery During Sleep." *Science*, Published Online 4 April 2013.

Jackson, Frank. (1982). "Epiphenomenal Qualia". *Philosophical Quarterly* 32, 127–136.

Kikkawa, M., Ishikawa, T., Nakata, T., Wakabayashi, T., and Hirokawa, N. (1994). "Direct visualization of the microtubule lattice seam both in vitro and in vivo". *Journal of Cell Biology* 127, no. 6, 1965–1971.

Leffingwell, John C. (1999). "Olfaction - A Review." Leffingwell & Associates. http://www.leffingwell.com/olfaction.htm, April 2013.

Leopold, D.A. and Logothetis, N.K. (1999). "Multistable phenomena: changing views in perception." *Trends in Cognitive Sciences* 3, no. 7, 254-264.

London, Fritz, and Bauer, Edmond. (1939). "La théorie de l'observation en mécanique quantique." English translation in *Quantum Theory and Measurement* 1983, 217-259.

Lutz, A., Greischar, L.L., Rawlings, N.B., Ricard, M., and Davidson, R.J. (2004). "Long-term meditators self-induce high-amplitude gamma synchrony during mental practice." *Proceedings of the National Academy of Sciences 101, no. 46,* 16369–16373

Nagel, Thomas. (2012). *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature Is Almost Certainly False.* New York: Oxford University Press.

Nagel, Thomas. (1974). "What Is It Like to Be a Bat?" *The Philosophical Review* 83, no. 4, 435-450.

Nishimoto, S., Vu, A.T., Naselaris, T., Benjamini, Y., Yu, B., and Gallant, J.L. (2011). "Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies." *Current Biology* 21, no. 19, 1641–1646.

Packard, M., and Teather, L. (1998). “Amygdala modulation of multiple memory systems: hippocampus and caudate-putamen.” *Neurobiology of Learning and Memory* 69, 163-203.

Palmer, John. (2003). "ESP in the Ganzfeld: Analysis of a Debate." *Journal of Consciousness Studies* 10, no. 6–7, 51–68.

Penman, Danny. (2008). "Could there be proof to the theory that we're ALL psychic?". *Daily Mail UK*, 28–29.

Penrose, Roger. (1989). *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. Oxford: Oxford University Press.

Persinger, M.A., Roll, W.G., Tiller, S.G., Koren, S.A., and Cook , C.M. (2002). "Remote viewing with the artist Ingo Swann: neuropsychological profile, electroencephalographic correlates, magnetic resonance imaging (MRI), and possible mechanisms." *Perceptual and Motor Skills* 94, no. 3 pt. 1, 927–949.

Pouget, A. and Sejnowski, T. (1997). “Spatial Transformations in the Parietal Cortex Using Basis Functions.” *Journal of Cognitive Neuroscience* 9, no. 2, 222-237.

Ramchandran, V.S. and Hirstein, William. (1998). “The perception of phantom limbs”. *Brain* 121, 1603–1630.

Schwender, D., Madler, C., Klasing, S., Peter, K., and Pöppel, E. (1994). “Anesthetic Control of 40-Hz Brain Activity and Implicit Memory.” *Consciousness and Cognition* 3, no. 2, 129–147.

Searle, John. (2007). *Freedom and Neurobiology: Reflections on Free Will, Language, and Political Power*. New York: Columbia University Press.

Shimojo, S., Kamitani, Y., and Nishida, S. (2001). "Afterimage of perceptually filled-in surface". *Science* 293, no. 5535, 1677–1680.

Spottiswoode, S.J.P. and May, E.C. (2003). "Skin Conductance Prestimulus Response: Analyses, Artifacts and a Pilot Study." *Journal of Scientific Exploration* 17, no. 4, 617–641.

Stevenson, Ian. (1993). "Birthmarks and Birth Defects Corresponding to Wounds on Deceased Persons." *Journal of Scientific Exploration* 7, no. 4, 403-410.

Tegmark, M. (2000). "Importance of quantum decoherence in brain processes". *Physical Review E* 61, no. 4, 4194–4206.

Thaheld, F.H. (2005). “Does consciousness really collapse the wavefunction? A possible objective biophysical resolution of the measurement problem.” *BioSystems* 81, 113-124.

Tononi, Giulio. (2008). “Consciousness as Integrated Information: a Provisional Manifesto.” The *Biological Bulletin 215,* 216-242.

Varela, Francisco J. and Shear, Jonathan. (1999). *The View from Within: First-person Approaches to the Study of Consciousness*. Thorverton (UK): Imprint Academic.

Wahlman, Jessica. (2013). “I’m Not There: Consciousness and Alternatives to Physicalism.” Awaiting Publication.

Walker, Evan Harris. (2000). *The Physics of Conciousness: Quantum Minds and the Meaning of Life*. Cambridge, MA: Perseus.

Weeks, Eric. (2013). Personal Communication.

Zeki, S., Perry, R.J., and Bartels, A. (2003). "The Processing of Kinetic Contours in the Brain." *Cerebral Cortex* 13, no. 2, 189-202.