

Review

Quantum physics in neuroscience and psychology: a neurophysical model of mind–brain interaction

Jeffrey M. Schwartz¹, Henry P. Stapp² and Mario Beauregard^{3,4,5,*}

¹*UCLA Neuropsychiatric Institute, 760 Westwood Plaza, NPI Los Angeles, CA 90024-1759, USA*

²*Theoretical Physics Mailstop 5104/50A Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720-8162, USA*

³*Département de Psychologie, Centre de Recherche en Neuropsychologie Expérimentale et Cognition (CERNEC),* ⁴*Département de Radiologie, and* ⁵*Centre de Recherche en Sciences Neurologiques (CRSN), Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Québec H3C 3J7, Canada*

Neuropsychological research on the neural basis of behaviour generally posits that brain mechanisms will ultimately suffice to explain all psychologically described phenomena. This assumption stems from the idea that the brain is made up entirely of material particles and fields, and that all causal mechanisms relevant to neuroscience can therefore be formulated solely in terms of properties of these elements. Thus, terms having intrinsic mentalistic and/or experiential content (e.g. ‘feeling’, ‘knowing’ and ‘effort’) are not included as primary causal factors. This theoretical restriction is motivated primarily by ideas about the natural world that have been known to be fundamentally incorrect for more than three-quarters of a century. Contemporary basic physical theory differs profoundly from classic physics on the important matter of how the consciousness of human agents enters into the structure of empirical phenomena. The new principles contradict the older idea that local mechanical processes alone can account for the structure of all observed empirical data. Contemporary physical theory brings directly and irreducibly into the overall causal structure certain psychologically described choices made by human agents about how they will act. This key development in basic physical theory is applicable to neuroscience, and it provides neuroscientists and psychologists with an alternative conceptual framework for describing neural processes. Indeed, owing to certain structural features of ion channels critical to synaptic function, contemporary physical theory must in principle be used when analysing human brain dynamics. The new framework, unlike its classic-physics-based predecessor, is erected directly upon, and is compatible with, the prevailing principles of physics. It is able to represent more adequately than classic concepts the neuroplastic mechanisms relevant to the growing number of empirical studies of the capacity of directed attention and mental effort to systematically alter brain function.

Keywords: mind; consciousness; brain; neuroscience; neuropsychology; quantum mechanics

The only acceptable point of view appears to be the one that recognizes *both* sides of reality—the quantitative and the qualitative, the physical and the psychical—as compatible with each other, and can embrace them simultaneously.

(Pauli 1955, p. 208)

1. INTRODUCTION

The introduction into neuroscience and neuropsychology of the extensive use of functional brain imaging technology has revealed, at the empirical level, an important causal role of directed attention in cerebral functioning. The identification of brain areas involved in a wide variety of information processing functions concerning learning, memory and various kinds of symbol manipulation has been the subject of extensive and intensive investigation (see Toga &

Mazziotta 2000). Neuroscientists consequently now have a reasonably good working knowledge of the role of a variety of brain areas in the processing of complex information. But, valuable as these empirical studies are, they provide only the data for, not the answer to, the critical question of the causal relationship between the aspects of empirical studies that are described in psychological terms and those that are described in neurophysiological terms. In most of the cases, investigators simply assume that measurable-in-principle properties of the brain are the only factors needed to explain eventually the processing of the psychologically described information that occurs in neuropsychological experiments. This privileging of physically describable brain mechanisms as the core, and indeed final, explanatory vehicle for the processing of every kind of psychologically described data, is the foundational assumption of almost all contemporary biologically based cognitive neuroscience.

* Author for correspondence (mario.beauregard@umontreal.ca).

It is becoming increasingly clear, however, that there is at least one type of information processing and manipulation that does not readily lend itself to explanations that assume that all final causes are subsumed within brain, or more generally, central nervous system mechanisms. The cases in question are those in which the conscious act of wilfully altering the mode by which experiential information is processed itself changes, in systematic ways, the cerebral mechanisms used. There is a growing recognition of the theoretical importance of applying experimental paradigms that use directed mental effort to produce systematic and predictable changes in brain function (e.g. Beauregard *et al.* 2001; Ochsner *et al.* 2002). These wilfully induced brain changes are generally accomplished through training in, and the applied use of, cognitive reattribution and the attentional re-contextualization of conscious experience. Furthermore, an accelerating number of studies in the neuroimaging literature significantly support the thesis that, again, with appropriate training and effort, people can systematically alter neural circuitry associated with a variety of mental and physical states that are frankly pathological (Schwartz *et al.* 1996; Schwartz 1998; Musso *et al.* 1999; Paquette *et al.* 2003). A recent review of this and the related neurological literature has coined the term ‘self-directed neuroplasticity’ to serve as a general description of the principle that focused training and effort can systematically alter cerebral function in a predictable and potentially therapeutic manner (Schwartz & Begley 2002).

From a theoretical perspective, perhaps the most important aspect of this line of research is the empirical support it provides for a new science-based way of conceptualizing the interface between mind/consciousness and brain. Until recently, virtually all attempts to understand the functional activity of the brain have been based, at least implicitly, on some principles of classic physics that have been known to be fundamentally false for three-quarters of a century. According to the classic conception of the world, all causal connections between observables are explainable in terms of mechanical interactions between material realities. But this restriction on modes of causation is not fully maintained by the currently applied principles of physics, which consequently offer an alternative conceptual foundation for the scientific description and modelling of the causal structure of self-directed neuroplasticity.

The advantages for neuroscience and neuropsychology of using the conceptual framework of *contemporary* physics, as opposed to that of *classic* physics, stem from five basic facts. First, terms such as ‘feeling’, ‘knowing’ and ‘effort’, because they are intrinsically mentalistic and experiential, cannot be described exclusively in terms of material structure. Second, to explain the observable properties of large physical systems that depend sensitively upon the behaviours of their atomic constituents, the founders of contemporary physical theory were led to introduce explicitly into *the basic causal structure of physics* certain important choices made by human beings about how they will act. Third, within this altered conceptual framework these choices are described in mentalistic

(i.e. psychological) language. Fourth, terminology of precisely this kind is critically necessary for the design and execution of the experiments in which the data demonstrating the core phenomena of self-directed neuroplasticity are acquired and described. Fifth, the injection of psychologically described choices on the part of human agents into the causal theoretical structure can be achieved for experiments in neuroscience by applying the same mathematical rules that were developed to account for the structure of phenomena in the realm of atomic science.

The consequence of these facts is that twentieth century physics, in contrast to classic physics, provides a rationally coherent pragmatic framework in which the psychologically and neurophysiologically described aspects of the neuroscience experiments mentioned above are causally related to each other in mathematically specified ways. Thus, contemporary physics allows the data from the rapidly emerging field of self-directed neuroplasticity to be described and understood in a way that is more rationally coherent, scientific and useful than what is permitted by theories in which all causation is required to be fundamentally mechanical.

To explicate the physics of the interface between mind/consciousness and the physical brain, we shall in this article describe in detail how the quantum mechanically based causal mechanisms work, and show why it is necessary in principle to advance to the quantum level to achieve an adequate theory of the neurophysiology of volitionally directed activity. The reason, essentially, is that classic physics is an approximation to the more accurate quantum theory, and that this classic approximation *eliminates the causal efficacy of our conscious efforts* that these experiments empirically manifest.

We shall also explain how certain structural features of ion conductance channels critical to synaptic function *entail* that the classic approximation fails in principle to cover the dynamics of a human brain. Quantum dynamics *must be used in principle*. Furthermore, once the transition to the quantum description is made, the principles of quantum theory must, in order to maintain rational consistency and coherency, be used to link the quantum physical description of the subject’s brain to their stream of conscious experiences. The conscious choices by human agents thereby become injected non-trivially into the causal interpretation of neuroscience and neuropsychology experiments. This caveat particularly applies to those experimental paradigms in which human subjects are required to perform decision-making or attention-focusing tasks that require conscious effort.

2. PRACTICAL RAMIFICATIONS OF THE ALTERED CONCEPT OF THE CAUSAL STRUCTURE OF SELF-DIRECTED NEUROPLASTICITY

Clarity is required about the sorts of neuroscientific reasoning that remain coherent, given the structure of modern physics and, contrastingly, the types of assertion that can now be viewed as the residue of a materialistic bias stemming from superseded physics. Entirely acceptable are *correlational analyses* about the

relationship between mentalistic data and neurophysiological mechanisms. Examining the qualitative and quantitative aspects of brain function, and doing detailed analyses of how they relate to the data of experience, obtained through increasingly sophisticated means of psychological investigation and subject self-report analysis (e.g. the entire September–October 2003 issue of *Journal of Consciousness Studies*, volume 10, number 9–10, is dedicated to these questions), are completely in line with fundamental physics. These activities are the core of neuropsychological science. What is not justified is the presumption, either tacit or explicit, that *all* aspects of experience examined and reported are necessarily causal consequences solely of brain mechanisms. The structure of contemporary physics entails no such conclusion. This is particularly relevant to data from first-person reports about active, wilfully directed attentional focus, and especially to data pertaining to which aspects of the stream of conscious awareness a subject chooses to focus on when making self-directed efforts to modify and/or modulate the quality and beam of attention. In such cases, the structure of orthodox quantum physics implies that the investigator is not justified in assuming that the focus of attention is determined wholly by brain mechanisms that are in principle completely well-defined and mechanically determined. Conscious effort itself can, justifiably within science, be taken to be a primary variable whose complete causal origins may be *untraceable in principle*, but whose causal efficacy in the physical world can be explained on the basis of the laws of physics.

As already emphasized, the cognitive frame in which neuroscience research, including research on cerebral aspects of behaviour, is generally conducted contains within it the assumption that brain mechanisms are in principle fully sufficient to explain all of the observed phenomena. In the fields of functional neuroimaging, this has led to experimental paradigms that focus primarily on changes in brain activation as primary variables used to explain whatever behavioural changes are observed—including ones understood as involving essentially cognitive and emotional responses. As long as one is investigating phenomena that are mostly passive in nature this may be fully justified. A person is shown a picture depicting an emotionally or perhaps a sexually arousing scene. The relevant limbic and/or diencephalic structures are activated. The investigator generally concludes that the observed brain activation has some intrinsic causal role in the emotional changes reported (or, perhaps, the hormonal correlates of those changes).

This method is all well and good, as far as it goes. In addition, from the experimental subject's perspective, it is all quite passive—all that is really required on his or her part is to remain reasonably awake and alert or, more precisely, at least somewhat responsive to sensory inputs. But when, as happens in a growing number of studies, the subject makes an active response aimed at systematically *altering* the nature of the emotional reaction—for example, by actively performing a cognitive reattribution—then the demand that the data be understood solely from the perspective of brain-based causal mechanism is a severe and counter-intuitive

constraint. It is noteworthy that this demand for an entirely brain-based causal mechanism is nullified, in the quantum model developed here, by a specified quantum effect, which will be described in detail below.

Surmounting the limitations imposed by restricting one's ideas to the failed concepts of classic physics can be especially important when one is investigating how to develop improved methods for altering the emotional and cerebral responses to significantly stressful external or internally generated stimuli. An incorrect assignment of the causal roles of neurophysiologically and mentalistically described variables can impact negatively on a therapist's selection of a course of treatment, on a patient's capacity to recover, and on a neuroscientist's design of clinically relevant research programmes.

In the analysis and development of clinical practices involving psychological treatments and their biological effects, the possession and use of a rationally coherent and physically allowable conception of the causal relationship between mind and brain (or, if one prefers, mentalistic and neurophysiological variables) is critical. If one simply accepts the standard presumption that all aspects of emotional response are passively determined by neurobiological mechanisms, then the theoretical development of genuinely effective self-directed psychological strategies that produce real neurobiological changes can be impeded by the fact that one is using a theory that excludes from the dynamics what logically can be, and in our model actually are, key causal elements, namely our wilful choices.

The clinician's attention is thus directed away from what can be in many cases, at the level of actual practice, a powerful determinant of action, namely the subject's psychologically (i.e. mentalistically) framed commitment to act or think in specific ways. The therapist tends to become locked into the view that the *psychological* treatment of ailments *caused by neurobiological impairments* is not a realistic goal.

There is already a wealth of data arguing against this view. For instance, work in the 1990s on patients with obsessive compulsive disorder demonstrated significant changes in caudate nucleus metabolism and the functional relationships of the orbitofrontal cortex–striatum–thalamus circuitry in patients who responded to a psychological treatment using cognitive reframing and attentional refocusing as key aspects of the therapeutic intervention (for review, see [Schwartz & Begley 2002](#)). More recently, work by Beauregard and colleagues ([Paquette et al. 2003](#)) has demonstrated systematic changes in the dorsolateral prefrontal cortex and parahippocampal gyrus after cognitive-behavioural therapy for phobia of spiders, with brain changes significantly related to both objective measurements and subjective reports of fear and aversion. There are now numerous reports on the effects of self-directed regulation of emotional response, via cognitive reframing and attentional re-contextualization mechanisms, on cerebral function (e.g. [Schwartz et al. 1996](#); [Beauregard et al. 2001](#); [Ochsner et al. 2002](#); [Lévesque et al. 2003](#); [Paquette et al. 2003](#)).

The brain area generally activated in all the studies done so far on the self-directed regulation of emotional response is the prefrontal cortex, a cortical region also

activated in studies of cerebral correlates of wilful mental activity, particularly those investigating self-initiated action and the act of attending to one's own actions (Spence & Frith 1999; Schwartz & Begley 2002). There is, however, one aspect of wilful mental activity that seems particularly critical to emotional self-regulation, and that seems to be the critical factor in its effective application—the factor of focused dispassionate self-observation that, in a rapidly growing number of clinical psychology studies, has come to be called ‘mindfulness’ or ‘mindful awareness’ (Segal *et al.* 2002).

The mental act of clear-minded introspection and observation, variously known as mindfulness, mindful awareness, bare attention, the impartial spectator, etc., is a well-described psychological phenomenon with a long and distinguished history in the description of human mental states (Nyanaponika 2000). The most systematic and extensive exposition is in the canonical texts of classic Buddhism preserved in the Pali language, a dialect of Sanskrit. Because of the critical importance of this type of close attentiveness in the practice of Buddhist meditation, some of its most refined descriptions in English are in texts concerned with meditative practice (although it is of critical importance to realize that the mindful mental state does not require any specific meditative practice to acquire, and is *certainly not* in any sense a ‘trance-like’ state).

One particularly well-established description, using the name ‘bare attention’, is as follows:

Bare Attention is the clear and single-minded awareness of what actually happens *to* us and *in* us at the successive moments of perception. It is called ‘Bare’ because it attends just to the bare facts of a perception as presented either through the five physical senses or through the mind...without reacting to them.

(Nyanaponika 1973, p. 30)

Perhaps the essential characteristic of mindful observation is that you are just watching, observing all facts, both inner and outer, very calmly, clearly and closely. To sustain this attentional perspective over time, especially during stressful events, invariably requires the conscious application of effort.

A working hypothesis for ongoing investigation in human neurophysiology, based on a significant body of preliminary data, is that the mental action of mindful awareness specifically modulates the activity of the prefrontal cortex. Because of the well-established role of this cortical area in the planning and wilful selection of self-initiated responses (Spence & Frith 1999; Schwartz & Begley 2002), the capacity of mindful awareness, and by implication all emotional self-regulating strategies, to specifically modulate activity in this critical brain region has tremendous implications for the fields of mental health and related areas.

It might be claimed that the designs and executions of successful clinical practices (and of informative neuropsychological experiments) that depend on the idea of the causal efficacy of conscious effort, and which fit so well into the quantum conceptualization that actually explains the causal efficacy of these efforts, could just as well be carried out within the conceptual framework in which the causal efficacy of wilful effort is

an illusion, or is something very different from what it intuitively seems to be. But such a claim is not easy to defend. Simple models that are consistent with basic intuition and lead directly to experimentally demonstrable conclusions are better than philosophically intricate ones that lead to the same conclusions. Of course, if it could be argued that the simple model could not be true because it violates the basic principles of physics whereas the more intricate one obeys them, then there might be reasonable grounds for question or dispute. But in the present case the reverse is true: it is the simple model that is built on the basic laws of physics and it is the arcane and philosophically difficult model, in which our basic human intuition concerning the efficacy of mental effort is denied as not being what it seems to be, which contradicts the laws of physics.

The major theoretical issue we address in this article is the failure of classic models of neurobiological action to provide a scientifically adequate account for all of the mechanisms that are operating when human beings use self-directed strategies for the purpose of modulating emotional responses and their cerebral correlates. Specifically, the assumption that all aspects of mental activity and emotional life are ultimately explicable solely in terms of micro-local deterministic brain activity, with no superposed effects of mental effort, produces a theoretical structure that both fails to meet practical scientific needs, and also fails to accord with the causal structure of modern physics.

In the alternative approach the role played by the mind, when one is observing and modulating one's own emotional states, is an intrinsically active and physically efficacious process in which mental action is affecting brain activity in a way concordant with the laws of physics. A culturally relevant way of framing this change is to say that contemporary physics imbues the venerable and therapeutically useful term ‘psychodynamic’ with rigorous neurophysical efficacy.

This new theory of the mind–brain connection is supportive of clinical practice. Belief in the efficacy of mental effort in emotional self-regulation is needed to subjectively access the phenomena (e.g. belief in the efficacy of effort is required to sustain mindfulness during stressful events). Moreover, a conceptual framework in which psychologically described efforts have effects is needed to explain to patients what they are supposed to do when directing their inner resources to the challenging task of modifying emotional and cerebral responses. Clinical success is jeopardized by a belief on the part of either therapists or patients that their mental effort is an illusion or a misconception.

It takes effort for people to achieve therapeutic results. That is because it requires a redirection of the brain's resources away from lower level limbic responses and toward higher level prefrontal functions—and this does not happen passively. Rather, it requires, in actual practice, both wilful training and directed mental effort. It is semantically inconsistent and clinically counterproductive to insist that these kinds of brain changes be viewed as being solely an intra-cerebral ‘the physical brain changing itself’ type of action. That is because practical aspects of the activity of mind essential to the identification, activation, application and use of directed mental effort are

not describable solely in terms of material brain mechanisms. The core phenomena necessary for the scientific description of self-directed neuroplasticity are processes that cannot be elaborated solely in terms of classic models of physics.

Furthermore, as we will see in detail in the following sections of this article, orthodox concepts of contemporary physics are ideally suited to a rational and practically useful understanding of the action of mindful self-observation on brain function. Classic models of physics, which view all action in the physical world as being ultimately the result of the movements of material particles, are now seriously out of date, and no longer need be seen as providing the unique, or the best, scientifically well-grounded paradigm for investigating the interface between mind/consciousness and brain.

When people practice self-directed activities for the purpose of systematically altering patterns of cerebral activation they are attending to their mental and emotional *experiences*, not merely their limbic or hypothalamic brain mechanisms. And although no scientifically oriented person denies that those brain mechanisms play a critical role in generating those experiences, precisely what the person is training him- or herself to do is to wilfully *change* how those brain mechanisms operate—and to do that requires attending to mental experience *per se*. It is, in fact, the basic thesis of self-directed neuroplasticity research that the *way in which a person directs their attention* (e.g. mindfully or unmindfully) will affect both the experiential state of the person and the state of his/her brain. The existence of this close connection between mental effort and brain activity flows naturally out of the dynamic principles of contemporary physics, but is, within the framework of classic physics, a difficult problem that philosophers of the mind have been intensively engaged with, particularly for the past 50 years. The core question is whether the solution to this problem lies wholly in the eventual development of a more sophisticated philosophy that is closely aligned with the classic known-to-be-fundamentally-false conception of nature, or whether the profound twentieth century development in physics, that assigns a subtle but essential causal role to human consciousness, can usefully inform our understanding of the effects of human consciousness in neuropsychological experiments that appear to exhibit the causally efficacious presence of such effects.

To appreciate the major conceptual changes made in basic physical theory during the twentieth century, one must know about certain key features of the older theory.

3. CLASSIC PHYSICS

Classic physics is a theory of nature that originated with the work of Isaac Newton in the seventeenth century and was advanced by the contributions of James Clerk Maxwell and Albert Einstein. Newton based his theory on the work of Johannes Kepler, who found that the planets appeared to move in accordance with a simple mathematical law, and in ways wholly determined by their spatial relationships to other objects. Those

motions were apparently *independent of our human observations of them*.

Newton effectively assumed that all physical objects were made of tiny miniaturized versions of the planets, which, like the planets, moved in accordance with simple mathematical laws, independently of whether we observed them or not. He found that he could then explain the motions of the planets and also the motions of large terrestrial objects and systems, such as cannon balls, falling apples and the tides, by assuming that every tiny planet-like particle in the solar system attracted every other one with a force inversely proportional to the square of the distance between them.

This force was an *instantaneous action at a distance*: it acted instantaneously, no matter how far the particles were apart. This feature troubled Newton. He wrote to a friend ‘That one body should act upon another through the vacuum, without the mediation of anything else, by and through which their action and force may be conveyed from one to another, is to me so great an absurdity that I believe no man, who has in philosophical matters a competent faculty of thinking, can ever fall into it’ (Newton 1687, p. 634). Although Newton’s philosophical persuasion on this point is clear, he nevertheless formulated his universal law of gravity without specifying how it was mediated.

Albert Einstein, building on the ideas of Maxwell, discovered a suitable mediating agent, a distortion of the structure of space–time itself. Einstein’s contributions made classic physics into what is called a *local theory*: there is no action at a distance. All influences are transmitted essentially by contact interactions between tiny neighbouring mathematically described ‘entities’, and no influence propagates faster than the speed of light.

Classic physics is, moreover, *deterministic*: the interactions are such that the state of the physical world at any time is completely determined by the state at any earlier time. Consequently, according to classic theory, the complete history of the physical world *for all time* is mechanically fixed by contact interactions between tiny component parts, together with the initial condition of the primordial universe.

This result means that, according to classic physics, *you are a mechanical automaton*: your every physical action was predetermined before you were born solely by mechanical interactions between tiny mindless entities. Your mental aspects are *causally redundant*: everything you do is completely determined by mechanical conditions alone, without any mention of your thoughts, ideas, feelings or intentions. Your intuitive feeling that your conscious intentions make a difference in what you do is, according to the principles of classic physics, a false and misleading illusion.

There are two possible ways within classic physics to understand this total incapacity of your mental side (i.e. your stream of conscious thoughts and feelings) to make any difference in what you do. The first way is to consider your thoughts, ideas and feelings to be epiphenomenal by-products of the activity of your brain. Your mental side is then a causally impotent sideshow that *is produced*, or *caused*, by your brain, but that *produces* no reciprocal action back upon your brain.

The second way is to contend that each of your conscious experiences—each of your thoughts, ideas, or feelings—is the *very same thing* as some pattern of motion of various tiny parts of your brain.

4. PROBLEMS WITH CLASSIC PHYSICS

William James (1890, p. 138) argued against the first possibility, epiphenomenal consciousness, by claiming that ‘*The particulars of the distribution of consciousness, so far as we know them, points to its being efficacious.*’ He noted that consciousness seems to be ‘an organ, superadded to the other organs which maintain the animal in its struggle for existence; and the presumption of course is that it helps him in some way in this struggle, just as they do. But it cannot help him without being in some way efficacious and influencing the course of his bodily history.’ James said that the study described in his book ‘will show us that consciousness is at all times primarily a *selecting agency*.’ It is present when choices must be made between different possible courses of action. He further mentioned that ‘It is to my mind quite inconceivable that consciousness should have *nothing to do* with a business to which it so faithfully attends’ (1890, p. 136).

If mental processes and consciousness have no effect upon the physical world, then what keeps a person’s mental world aligned with their physical situation? What keeps their pleasures in general alignment with actions that benefit them, and pains in general correspondence with things that damage them, if felt pleasures and pains have no effect at all upon their actions?

These liabilities of the notion of epiphenomenal mind and consciousness lead many thinkers to turn to the alternative possibility that a person’s mind and stream of consciousness is *the very same thing* as some activity in their brain: mind and consciousness are ‘emergent properties’ of brains.

A huge philosophical literature has developed arguing for and against this idea. The primary argument against this ‘emergent-identity theory’ position, *within a classic physics framework*, is that in classic physics the full description of nature is in terms of numbers assigned to tiny space–time regions, and there appears to be no way to understand or explain how to get from such a restricted conceptual structure, which involves such a small part of the world of experience, to the whole. How and why should that extremely limited conceptual structure (which arose basically from idealizing, by miniaturization, certain features of observed planetary motions) suffice to explain the totality of experience, with its pains, sorrows, hopes, colours, smells and moral judgements? Why, *given the known failure of classic physics at the fundamental level*, should that richly endowed whole be explainable in terms of such a narrowly restricted part?

The core ideas of the arguments in favour of an identity-emergent theory of mind and consciousness are illustrated by Roger Sperry’s (1992) example of a ‘wheel’. A wheel obviously does something: it is causally efficacious; it carries the cart. It is also an *emergent property*: there is no mention of ‘wheelness’ in the formulation of the laws of physics and ‘wheelness’ did not exist in the early universe; ‘wheelness’ *emerges*

only under certain special conditions. And the macroscopic wheel exercises ‘top-down’ control of its tiny parts. All these properties are perfectly in line with classic physics, and with the idea that ‘a wheel is, precisely, a structure constructed out of its tiny atomic parts’. So why not suppose mind and consciousness to be, like ‘wheelness’, emergent properties of their classically conceived tiny physical parts?

The reason that mind and consciousness are not analogous to ‘wheelness’, within the context of classic physics, is that the properties that characterize ‘wheelness’ are properties that are *entailed*, within the conceptual framework of classic physics, by properties specified in classic physics, whereas the properties that characterize conscious mental processes, namely the various ways these processes feel, are not *entailed* within the conceptual structure provided by classic physics, but by the properties specified by classic physics.

That is the huge difference-in-principle that distinguishes mind and consciousness from things that, according to classic physics, are constructible out of the particles that are postulated to exist by classic physics.

Given the state of motion of each of the tiny physical parts of a wheel, as it is conceived of in classic physics, the properties that characterize the wheel (e.g. its roundness, radius, centre point, rate of rotation, etc.) are specified within the conceptual framework provided by the principles of classic physics, which specify only geometric-type properties such as changing locations and shapes of conglomerations of particles and numbers assigned to points in space. But given the state of motion of each tiny part of the brain, as it is conceived of in classic physics, the properties that characterize the stream of consciousness (the painfulness of the pain, the feeling of the anguish, or of the sorrow, or of the joy) are not specified, within the conceptual framework provided by the principles of classic physics. Thus it is possible, within that classic physics framework, to strip away those feelings without disturbing the physical descriptions of the motions of the tiny parts. One can, within the conceptual framework of classic physics, take away the consciousness while leaving intact the properties that enter into that theoretical construct, namely the locations and motions of the tiny physical parts of the brain and its physical environment. But one cannot, within the conceptual framework provided by classic physics, take away the physical characteristics that define the ‘wheelness’ of a wheel without affecting the locations and motions of the tiny physical parts of the wheel.

Because one can, within the conceptual framework provided by classic physics, strip away mind and consciousness without affecting the physical behaviour, one cannot rationally claim, *within that framework*, that mind and consciousness are the *causes* of the physical behaviour, or are *causally efficacious* in the physical world. Thus the ‘identity theory’ or ‘emergent property’ strategy fails in its attempt to make mind and consciousness efficacious, insofar as one remains strictly within the conceptual framework provided by classic physics. Moreover, the whole endeavour to base brain theory on classic physics is undermined by the fact that classic theory is unable to account for behavioural properties (such as electrical and thermal

conductivity, and elasticity, etc.) that depend sensitively upon the behaviour of the atomic, molecular and ionic constituents of a system, and brains are certainly systems of this kind, as will be discussed in detail later.

Although classic physics is unable to account for observable properties that depend sensitively on the behaviours of atoms, molecules and ions, the classic theory is an approximation to a more accurate theory, called quantum theory, which *is* able to account for these observable macroscopic properties. But if classic physics is unable to account for the moderately complex behavioural properties of most other large systems, then how can it be expected to account for the exquisitely complex behavioural properties of thinking brains?

5. THE QUANTUM APPROACH

Early in the twentieth century scientists discovered empirically that the principles of classic physics could not be correct. Moreover, those principles were wrong in ways that no minor tinkering could ever fix. The *basic principles* of classic physics were thus replaced by *new basic principles* that account uniformly for all the successes of the older classic theory and for all the data that are incompatible with the classic principles.

The key philosophical and scientific achievement of the founders of quantum theory was to forge a rationally coherent and practicable linkage between the two kinds of description that jointly comprise the foundation of science. Descriptions of the first kind are accounts of psychologically experienced empirical findings, expressed in a language that allows us to communicate to our colleagues what we have done and what we have learned. Descriptions of the second kind are specifications of physical properties, which are expressed by assigning mathematical properties to space–time points and formulating laws that determine how these properties evolve over the course of time. Bohr, Heisenberg, Pauli and the other inventors of quantum theory discovered a useful way to connect these two kinds of description by causal laws. Their seminal discovery was extended by John von Neumann from the domain of atomic science to the realm of neuroscience and, in particular, to the problem of understanding and describing the causal connections between the minds and the brains of human beings.

In order to achieve this result, the whole *concept of what science is* was turned inside out. The core idea of classic physics was to describe the ‘world out there’, with no reference to ‘our thoughts in here’. But the core idea of quantum mechanics is to describe both *our activities as knowledge-seeking and knowledge-acquiring agents*, and *the knowledge that we thereby acquire*. Thus, quantum theory involves, essentially, what is ‘in here’, not just what is ‘out there’.

This philosophical shift arises from the *explicit* recognition by quantum physicists that science is about *what we can know*. It is fine to have a beautiful and elegant mathematical theory about a *really existing physical world out there* that meets various intellectually satisfying criteria. But the essential demand of *science* is that the theoretical constructs be tied to the experiences of the human scientists who devise ways of testing

the theory and of the human engineers and technicians who both participate in these tests and eventually put the theory to work. Thus, the structure of a proper physical theory must involve not only the part describing the behaviour of the not-directly experienced theoretically postulated entities, expressed in some appropriate symbolic language, but also a part describing the human experiences that are pertinent to these tests and applications, expressed in the language that we actually use to describe such experiences to ourselves and to each other. And the theory must specify the connection between these two differently described and differently conceived parts of scientific practice.

Classic physics meets this final requirement in a trivial way. The relevant experiences of the human participants are taken to be direct apprehensions of the gross properties of large objects composed of huge numbers of their tiny atomic-scale parts. These apprehensions (of, for example, the perceived location and motion of a falling apple or the position of a pointer on a measuring device) were taken to be *passive*: they had no effect on the behaviours of the systems being studied. But the physicists who were examining the behaviours of systems that depend sensitively upon the behaviours of their tiny atomic-scale components found themselves forced to introduce a less trivial theoretical arrangement. In the new scheme the human agents are no longer passive observers. They are considered to be *active agents or participants*.

The participation of the agent continues to be important *even when the only features of the physically described world being observed are large-scale properties of measuring devices*. The sensitivity of the behaviour of the devices to the behaviour of some tiny atomic-scale particles propagates first to the devices and then to the observers in such a way that the choice made by an observer about what sort of knowledge to seek can profoundly affect the knowledge that can ever be received either by that observer himself or by any other observer with whom he can communicate. Thus the *choice* made by the observer about how he or she will act at a macroscopic level has, at the practical level, a profound effect on the physical system being acted upon.

That conclusion is not surprising. How one acts on a system would, in general, be *expected* to affect it. Nor is it shocking that the effect of the agent’s actions upon the system being probed is specified by the quantum mechanical rules. But the essential point not to be overlooked is that the logical structure of the basic physical theory has become fundamentally transformed. The agent’s choice about how to act has been introduced into the scientific description at a basic level and in a way that specifies, mathematically, how his or her choice about how to act affects the physical system being acted upon.

The structure of quantum mechanics is such that, although the effect upon the observed system of the agent’s choice about how to act is mathematically specified, the manner in which this choice itself is determined is not specified. This means that, in the treatment of experimental data, the choices made by human agents must be treated as freely chosen input variables, rather than as mechanical consequences of

any known laws of nature. Quantum theory thereby converts science's concept of us from that of a mechanical automaton, whose conscious choices are mere cogs in a gigantic mechanical machine, to that of agents whose conscious free choices affect the physically described world in a way specified by the theory. The approximation that reduces quantum theory to classic physics completely eliminates the important element of conscious free choice. Hence, from a physics point of view, trying to understand the connection between mind/consciousness and brain by going to the classic approximation is absurd: it amounts to trying to understand something in an approximation that eliminates the effect we are trying to study.

This original formulation of quantum theory was created primarily at an institute in Copenhagen directed by Niels Bohr and is called 'the Copenhagen interpretation'. Owing to the strangeness of the properties of nature entailed by the new mathematics, the Copenhagen strategy was to refrain from making any ordinary sort of ontological claims, but instead to take an essentially pragmatic stance. Thus, the theory was formulated *basically* as a set of practical rules for how scientists should go about the tasks of acquiring, manipulating and using knowledge. Claims about 'what the world out there is really like' were considered to lie beyond *science*.

This change in perspective is captured by Heisenberg's famous statement:

The conception of the objective reality of the elementary particles has thus evaporated not into the cloud of some obscure new reality concept, but into the transparent clarity of a mathematics that represents no longer the behavior of the particle but rather our knowledge of this behavior.

(Heisenberg 1958, p. 100).

A closely connected change is encapsulated in Niels Bohr's dictum that 'in the great drama of existence we ourselves are both actors and spectators' (Bohr 1963, p. 15; 1958, p. 81). The emphasis here is on 'actors': in classic physics we were mere spectators. The key idea is more concretely expressed in statements such as:

The freedom of experimentation, presupposed in classic physics, is of course retained and corresponds to the free choice of experimental arrangement for which the mathematical structure of the quantum mechanical formalism offers the appropriate latitude.

(Bohr 1958, p. 73)

Copenhagen quantum theory is about how the choices made by conscious human agents affect the knowledge they can and do acquire about the physically described systems upon which these agents act. In order to achieve this re-conceptualization of physics the Copenhagen formulation separates the physical universe into two parts, which are described in two different languages. One part is the observing human agent plus its measuring devices. This extended 'agent', which includes the devices, is described in mental terms—in terms of our instructions to colleagues about how to set up the devices and our reports of what we then 'see', or otherwise consciously experience. The other part of nature is *the system that the agent is acting*

upon. That part is described in physical terms—in terms of mathematical properties assigned to tiny space–time regions. Thus, Copenhagen quantum theory brings 'doing science' into science. In particular, it brings a crucial part of doing science, namely our choices about how we will probe nature, directly into the causal structure. It specifies the effects of these probing actions upon the systems being probed.

This approach works very well in practice. However, the body and brain of the human agent, and also their devices, are composed of atomic constituents. Hence a complete theory ought to be able to describe these systems in physical terms.

The great mathematician and logician John von Neumann formulated quantum theory in a rigorous way that allows the bodies and brains of the agents, along with their measuring devices, to be shifted into the physically described world. This shift is carried out in a series of steps, each of which moves more of what the Copenhagen approach took to be the psychologically described 'observing system' into the physically described 'observed system'. At each step the crucial act of choosing or deciding between possible optional observing actions remains undetermined by the physical observed system. This act of choosing is always ascribed to the observing agent. In the end all that is left of this agent is what von Neumann calls his 'abstract ego'. It is described in psychological terms, and is, in practice, the stream of consciousness of the agent.

At each step the direct effect of the conscious act is upon the part of the physically described world that is closest to the psychologically described world. This means that, in the end, the causal effect of the agent's mental action is on their own brain, or some significant part of their brain.

von Neumann makes the logical structure of quantum theory very clear by identifying two very different processes, which he calls process 1 and process 2 (von Neumann 1955, p. 418). Process 2 is the analogue in quantum theory of the process in classic physics that takes the state of a system at one time to its state at a later time. This process 2, like its classic analogue, is *local* and *deterministic*. However, process 2 by itself is not the whole story: it generates a host of 'physical worlds', most of which do not agree with our human experience. For example, if process 2 were, from the time of the big bang, the *only* process in nature, then the quantum state (centre point) of the moon would represent a structure smeared out over a large part of the sky, and each human body–brain would likewise be represented by a structure smeared out continuously over a huge region. Process 2 generates a *cloud of possible worlds*, instead of the one world we actually experience.

This huge disparity between properties generated by the 'mechanical' process 2 and the properties we actually observe is resolved by invoking process 1.

Any physical theory must, in order to be complete, specify how the elements of the theory are connected to human experience. In classic physics this connection is part of a *metaphysical* superstructure: it is not part of the dynamic process. But in quantum theory a linkage of the mathematically described physical state to human experiences is contained in the mathematically

specified dynamic. This connection is not passive. It is not a mere *witnessing* of a physical feature of nature. Instead, *it injects into the physical state of the system being acted upon specific properties that depend upon choices made by the agent.*

Quantum theory is built upon the practical concept of intentional actions by agents. Each such action is a *preparation* that is expected or intended to produce an experiential response or feedback. For example, a scientist might act to place a Geiger counter near a radioactive source and expect to see the counter either ‘fire’ during a certain time-interval or not ‘fire’ during that interval. The experienced response, ‘Yes’ or ‘No’, to the question, ‘Does the counter fire during the specified interval?’, specifies one bit of information. Quantum theory is thus an information-based theory built upon the preparative actions of information-seeking agents.

Probing actions of this kind are not only performed by scientists. Every healthy and alert infant is continually engaged in making wilful efforts that produce experiential feedbacks and he or she soon begins to form expectations about what sorts of feedbacks are probable to follow from some particular kind of effort. Thus, both empirical science and normal human life are based on paired realities of this action–response kind, and our physical and psychological theories are both basically attempting to understand these linked realities within a rational conceptual framework.

The basic building blocks of quantum theory are, then, a set of intentional actions by agents and for each such action an associated collection of possible ‘Yes’ feedbacks, which are the possible responses that the agent can judge to be in conformity to the criteria associated with that intentional act. For example, the agent is assumed to be able to make the judgement ‘Yes’ the Geiger counter clicked, or ‘No’ the Geiger counter did not click. Science would be difficult to pursue if scientists could make no such judgements about what they are experiencing.

All known physical theories involve idealizations of one kind or another. In quantum theory the main idealization is not that every object is made up of miniature planet-like objects. It is rather that there are agents that perform intentional acts each of which can result in feedback that may or may not conform to a certain criterion associated with that act. One piece of information is introduced into the world in which that agent lives, according to whether or not the feedback conforms to that criterion. The answer places the agent on one or the other of two alternative possible branches of the course of world history.

These remarks reveal the enormous difference between classic physics and quantum physics. In classic physics the elemental ingredients are tiny invisible bits of matter that are idealized miniaturized versions of the planets that we see in the heavens and that move in ways unaffected by our scrutiny, whereas in quantum physics the elemental ingredients are intentional preparative actions by agents, the feedbacks arising from these actions and the effects of these actions upon the physically described states of the probed systems.

This radical restructuring of the form of physical theory grew out of a seminal discovery by Heisenberg.

That discovery was that in order to get a satisfactory quantum generalization of a classic theory one must replace various *numbers* in the classic theory by *actions* (operators). A key difference between numbers and actions is that if A and B are two actions then AB represents the action obtained by performing the action A upon the action B. If A and B are two different actions then generally AB is different from BA: the order in which actions are performed matters. But for numbers the order does not matter: AB=BA.

The difference between quantum physics and its classic approximation resides in the fact that in the quantum case certain differences AB–BA are proportional to a number measured by Max Planck in 1900, and called Planck’s constant. Setting those differences to zero gives the classic approximation. Thus quantum theory is closely connected to classic physics, but is incompatible with it, because certain *non-zero* quantities must be replaced by zero to obtain the classic approximation.

The intentional actions of agents are represented mathematically in Heisenberg’s space of actions. A description of how it operates follows.

Each intentional action depends, of course, on the *intention of the agent* and upon the *state of the system* upon which this action acts. Each of these *two aspects of nature* is represented within Heisenberg’s space of actions by an *action*. The idea that a ‘state’ should be represented by an ‘action’ may sound odd, but Heisenberg’s key idea was to replace what classic physics took to be a ‘being’ with a ‘doing’. We shall denote the *action* (or operator) that represents the state being acted upon by the symbol *S*.

An intentional act is an action that is intended to produce a feedback of a certain conceived or imagined kind. Of course, no intentional act is certain: one’s intentions may not be fulfilled. Hence the intentional action merely puts into play a process that will lead either to a confirmatory feedback ‘Yes’, the intention is realized, or to the result ‘No’, the ‘Yes’ response did not occur.

The effect of this intentional mental act is represented mathematically by an equation that is one of the key components of quantum theory. This equation represents, within quantum mathematics, the effect of process 1 action upon the quantum state *S* of the system being acted upon. The equation is:

$$S \rightarrow S' = PSP + (I - P)S(I - P).$$

This formula exhibits the important fact that this process 1 action changes the state *S* of the system being acted upon into a new state *S'*, which is a sum of two parts.

The first part, *PSP*, represents in physical terms, the possibility in which the experiential feedback called ‘Yes’ appears and the second part, *(I–P)S(I–P)*, represents the alternative possibility ‘No’, this ‘Yes’ feedback does not appear. Thus, an effect of the probing action is injected into the mathematical description of the physical system being acted upon.

The operator *P* is important. The action represented by *P*, acting both on the right and on the left of *S*, is the

action of eliminating from the state S all parts of S except the ‘Yes’ part. That particular retained part is determined by the choice made by the agent. The symbol I is the unit operator, which is essentially multiplication by the number 1, and the action of $(I - P)$, acting both on the right and on the left of S is, analogously, to eliminate from S all parts of S except the ‘No’ parts.

Notice that process 1 produces the *sum* of the two alternative possible feedbacks, not just one or the other. Since the feedback must either be ‘Yes’ or ‘No=Not-Yes’, one might think that process 1, which keeps both the ‘Yes’ and the ‘No’ possibilities, would do nothing. But that is not correct. This is a key point. It can be made absolutely clear by noticing that S can be written as a sum of four parts, only two of which survive the process 1 action:

$$S = PSP + (I - P)S(I - P) + PS(I - P) + (I - P)SP.$$

This formula is a strict identity. The dedicated reader can quickly verify it by collecting the contributions of the four occurring terms PSP , PS , SP and S , and verifying that all terms but S cancel out. This identity shows that the state S is a sum of four parts, *two of which are eliminated by process 1*.

But this means that process 1 has a non-trivial *effect* upon the state being acted upon: it eliminates the two terms that correspond neither to the appearance of a ‘Yes’ feedback nor to the failure of the ‘Yes’ feedback to appear.

This result is the *first key point*: quantum theory has a specific causal process, process 1, which produces a non-trivial effect of an *agent’s choice* upon the physical description of the system being examined. (‘Nature’ will eventually choose between ‘Yes’ and ‘No’, but we focus here on the prior process 1, *the agent’s choice*. Nature’s subsequent choice we shall call process 3.)

(a) *Free choices*

The second key point is this: the agent’s choices are ‘free choices’, *in the specific sense specified below*.

Orthodox quantum theory is formulated in a realistic and practical way. It is structured around the activities of human agents, who are considered able to freely elect to probe nature in any one of many possible ways. Bohr emphasized the freedom of the experimenters in passages such as the one already quoted earlier, or the similar:

The foundation of the description of the experimental conditions as well as our freedom to choose them is fully retained.

(Bohr 1958, p. 90)

This freedom of choice stems from the fact that in the original Copenhagen formulation of quantum theory the human experimenter is considered to stand outside the system to which the quantum laws are applied. Those quantum laws are the only precise laws of nature recognized by that theory. Thus, according to the Copenhagen philosophy, *there are no presently known laws that govern the choices* made by the agent/experimenter/observer about how the observed system is to be probed. This choice is thus, *in this very*

specific sense, a ‘free choice’. The von Neumann generalization leaves this freedom intact. The choices attributed to von Neumann’s ‘abstract ego’ are no more limited by the known rules of quantum theory than are the choices made by Bohr’s experimenter.

(b) *Nerve terminals, ion channels and the need to use quantum theory in the study of the mind–brain connection*

Neuroscientists studying the connection of mind and consciousness to physical processes in the brain often assume that a conception of nature based on classic physics will eventually turn out to be adequate. That assumption would have been reasonable during the nineteenth century. But now, in the twenty-first century, it is rationally untenable. Quantum theory must be used in principle because the behaviour of the brain depends sensitively upon atomic, molecular and ionic processes, and these processes in the brain often involve large quantum effects.

To study quantum effects in brains within an orthodox (i.e. Copenhagen or von Neumann) quantum theory one must use the von Neumann formulation. This is because *Copenhagen* quantum theory is formulated in a way that leaves out the quantum dynamics of the human observer’s body and brain. But von Neumann quantum theory takes the physical system S upon which the crucial process 1 acts to be precisely the brain of the agent, or some part of it. Thus process 1 describes here an interaction between a person’s stream of consciousness, described in mentalistic terms, and an activity in their brain, described in physical terms.

A key question is the *quantitative magnitude* of quantum effects in the brain. They must be large in order for deviations from classic physics to play any significant role. To examine this quantitative question we consider the quantum dynamics of nerve terminals.

Nerve terminals are essential connecting links between nerve cells. The general way they work is reasonably well understood. When an action potential travelling along a nerve fibre reaches a nerve terminal, a host of ion channels open. Calcium ions enter through these channels into the interior of the terminal. These ions migrate from the channel exits to release sites on vesicles containing neurotransmitter molecules. A triggering effect of the calcium ions causes these contents to be dumped into the synaptic cleft that separates this terminal from a neighbouring neuron, and these neurotransmitter molecules influence the tendencies of that neighbouring neuron to ‘fire’.

At their narrowest points, calcium ion channels are less than a nanometre in diameter (Cataldi *et al.* 2002). This extreme smallness of the opening in the calcium ion channels has profound quantum mechanical implications. The narrowness of the channel restricts the lateral *spatial* dimension. Consequently, the lateral *velocity* is forced by the *quantum uncertainty principle* to become large. This causes the *quantum cloud of possibilities* associated with the calcium ion to *fan out* over an increasing area as it moves away from the tiny channel to the target region where the ion will be absorbed as a whole, or not absorbed at all, on some small triggering site.

This spreading of this *ion wave packet* means that the ion may or may not be absorbed on the small triggering site. Accordingly, the contents of the vesicle may or may not be released. Consequently, the quantum state of the brain has a part in which the neurotransmitter is released and a part in which the neurotransmitter is not released. This quantum splitting occurs at every one of the trillions of nerve terminals. This means that the quantum state of the brain splits into a vast host of classically conceived possibilities, one for each possible combination of the release-or-no-release options at each of the nerve terminals. In fact, because of uncertainties on timings and locations, what is generated by the physical processes in the brain will be not a single discrete set of non-overlapping physical possibilities but rather a huge *smear* of classically conceived possibilities. Once the physical state of the brain has evolved into this huge smear of possibilities one must appeal to the quantum rules, and in particular to the effects of process 1, in order to connect the physically described world to the streams of consciousness of the observer/participants.

This focus on the motions of calcium ions in nerve terminals is not meant to suggest that this particular effect is the *only* place where quantum effects enter into the brain process, or that the quantum process 1 acts locally at these sites. What is needed here is only the existence of *some* large quantum of effect. The focus upon these calcium ions stems from the facts that (i) in this case the various sizes (dimensions) needed to estimate the *magnitude* of the quantum effects are empirically known, and (ii) that the release of neurotransmitter into synaptic clefts is known to play a significant role in brain dynamics.

The brain matter is warm and wet and is continually interacting intensely with its environment. It might be thought that the strong *quantum decoherence effects* associated with these conditions would wash out all quantum effects, beyond localized chemical processes that can be conceived to be imbedded in an essentially classic world.

Strong decoherence effects are certainly present, but they are automatically taken into account in the von Neumann formulation employed here. These effects merely convert the state S of the brain into what is called a ‘statistical mixture’ of ‘nearly classically describable’ states, each of which develops in time (in the absence of process 1 events), in an almost classically describable way.

The existence of strong decoherence effects makes the main consequences of quantum theory being discussed here more easily accessible to neuroscientists by effectively reducing the complex quantum state of the brain to a collection of almost classically describable possibilities. Because of the uncertainties introduced at the ionic, atomic, molecular and electronic levels, the brain state will develop not into one single classically describable macroscopic state, as it does in classic physics, but into a continuous distribution of parallel virtual states of this kind. Process 1 must then be invoked to allow definite empirical predictions to be extracted from this continuous smear of parallel overlapping almost-classic possibilities generated by process 2.

(c) *Quantum brain dynamics*

A principal function of the brain is to receive clues from the environment, to form an appropriate plan of action and to direct and monitor the activities of the brain and body specified by the selected plan of action. The exact details of the plan will, for a classic model, obviously depend upon the exact values of many noisy and uncontrolled variables. In cases close to a bifurcation point, the dynamic effects of noise might even tip the balance between two very different responses to the given clues, for example, tip the balance between the ‘fight’ or ‘flight’ response to some shadowy form. It is important to realize that the exact values accounting for what in classic physics models are called ‘dynamic effects of noise’ are unknowable in principle. The contemporary physical model accounts for these uncertainties in brain dynamics.

The effect of the independent ‘release’ or ‘do not release’ options at each of the trigger sites, coupled with the uncertainty in the timing of the vesicle release at each of the trillions of nerve terminals, will be to cause the quantum mechanical state of the brain to become a smeared-out cloud of different macroscopic possibilities, some representing different alternative possible plans of action. As long as the brain dynamic is controlled wholly by process 2—which is the quantum generalization of the Newtonian laws of motion of classic physics—all of the various alternative possible plans of action will exist in parallel, with no one plan of action singled out as the one that will actually be experienced.

Some process beyond the local deterministic process 2 is required to pick out one experienced course of physical events from the smeared-out mass of possibilities generated by all of the alternative possible combinations of vesicle releases at all of the trillions of nerve terminals. As already emphasized, this other process is process 1. This process brings in a *choice* that is not determined by any currently known law of nature, yet has a definite effect upon the brain of the chooser. The process 1 choice picks an operator P and also a time t at which P acts. The effect of this action at time t is to change the state $S(t)$ of the brain, or of some large part of the brain, to

$$PS(t)P + (I - P)S(t)(I - P).$$

The action P cannot act at a *point* in the brain, because action at a point would dump a huge (in principle infinite) amount of energy into the brain, which would then explode. The operator P *must*, therefore, act non-locally, over a potentially large part of the brain.

In examining the question of the nature of the effect in the brain of process 2 we focused on the separate motions of the individual particles. But the physical structures in terms of which the action of process 1 is naturally expressed are not the separate motions of individual particles. They are, rather, the *quasi-stable macroscopic degrees of freedom*. The brain structures selected by the action of P must enjoy the stability, endurance and causal linkages needed to bring the intended experiential feedbacks into being.

These functional structures are probably more like the lowest-energy state of the simple harmonic

oscillator, which is completely stable, or like the states obtained from such lowest-energy states by spatial displacements and shifts in velocity. These shifted states tend to *endure as oscillating states*. In other words, in order to create the needed causal structure the projection operator P corresponding to an intentional action ought to pick out functionally pertinent quasi-stable *oscillating states of macroscopic subsystems of the brain*. The state associated with a process 1 preparatory intervention should be a functionally important brain analogue of a collection of oscillating modes of a drumhead, in which large assemblies of particles are moving in a coordinated way. Such an enduring structure in the brain can serve as a trigger and coordinator of further coordinated activities.

(d) *Templates for action*

The brain process that is actualized by the transition $S(t) \rightarrow PS(t)P$ is the neural correlate of the psychologically intended action. It is the brain's template for the intended action. It is a pattern of neuroelectrical activity that, if held in place long enough, will tend to generate a physical action in the brain that will tend to produce the intended experiential feedback.

(e) *Origin of the choices of the process 1 actions*

It has been repeatedly emphasized here that the choices by which process 1 actions actually occur are 'free choices' in the sense that they are not specified by the currently known laws of physics. On the other hand, a person's intentions are surely related in some way to their historical past. This means that the laws of contemporary orthodox quantum theory, although restrictive and important, do not provide a complete picture. In spite of this, orthodox quantum theory, while making no claim to ontological completeness, is able to achieve a certain kind of *pragmatic completeness*. It does so by treating the process 1 'free choices' as the input variables of experimental protocols, rather than mechanically determined consequences of brain action.

In quantum physics the 'free choices' made by human subjects are regarded as *subjectively controllable input variables*. Bohr emphasized that 'the mathematical structure of the quantum mechanical formalism offers the appropriate latitude' for these free choices. But the need for this strategic move goes deeper than the mere fact that contemporary quantum theory fails to specify how these choices are made. For if in the von Neumann formulation one does seek to determine the cause of the 'free choice' within the representation of the physical brain of the chooser, one finds that one is *systematically* blocked from determining the cause of the choice by the Heisenberg uncertainty principle, which asserts that the locations and velocities of, say, the calcium ions, are simultaneously unknowable to the precision needed to determine what the choice will be. Thus, one is not only faced with merely a practical unknowability of the causal origin of the 'free choices', but with an *unknowability in principle* that stems from the uncertainty principle itself, which lies at the base of quantum mechanics. There is thus a deep root in quantum theory for the idea that the origin of the 'free choices' does not lie in the physical description alone and also

for the consequent policy of treating these 'free choices' as empirical inputs that are selected by agents and enter into the causal structure via process 1.

(f) *Effort*

It is useful to classify process 1 events as either 'active' or 'passive'. *Passive* process 1 events are attentional events that occur with little or no feeling of conscious effort. *Active* process 1 events are intentional and involve effort. This distinction is given a functional significance by allowing 'effort' to enter into the selection of process 1 events in a way that will now be specified.

Consciousness probably contributes very little to brain dynamics compared with the contribution of the brain itself. To minimize the input of consciousness, and in order to achieve testability, we propose to allow mental effort to do nothing but control 'attention density', which is the rapidity of the process 1 events. This allows effort to have only a very limited kind of influence on brain activities, which are largely controlled by physical properties of the brain itself.

The notion that only the attention density is controlled by conscious effort arose from an investigation into what sort of conscious control over process 1 action was sufficient to accommodate the most blatant empirical facts. Imposing this strong restriction on the allowed effects of consciousness produces a theory with correspondingly strong predictive power. In this model all significant effects of consciousness upon brain activity arise exclusively from a well-known and well-verified strictly quantum effect known as the 'quantum Zeno effect' (QZE).

(g) *The quantum Zeno effect*

If one considers only passive events, then it is very difficult to identify any empirical effect of process 1, apart from the occurrence of awareness. In the first place, the empirical averaging over the 'Yes' and 'No' possibilities in strict accordance with the quantum laws tends to wash out all effects that depart from what would arise from a classic statistical analysis that incorporates the uncertainty principle as simply lack of knowledge. Moreover, the passivity of the mental process means that we have no empirically controllable variable.

However, the study of effortfully controlled intentional action brings in two empirically accessible variables, the intention and the amount of effort. It also brings in the important physical QZE. This effect is named for the Greek philosopher Zeno of Elea, and was brought into prominence in 1977 by the physicists Misra & Sudarshan (1977). It gives a name to the fact that repeated and closely spaced observational acts can effectively hold the 'Yes' feedback in place for an extended time-interval that depends upon the *rapidity at which the process 1 actions are happening*. According to our model, this rapidity is controlled by the amount of effort being applied. In our notation, the effect is to keep the 'Yes' condition associated with states of the form PSP in place longer than would be the case if no effort were being made. This 'holding' effect can override very strong mechanical forces arising from process 2.

The ‘Yes’ states *PSP* are assumed to be conditioned by training and learning to contain the template for action which if held in place for an extended period will tend to produce the intended experiential feedback. Thus, the model allows intentional mental efforts to tend to bring intended experiences into being. Systems that have the capacity to exploit this feature of natural law, as it is represented in quantum theory, would apparently enjoy a tremendous survival advantage over systems that do not or cannot exploit it.

6. SUPPORT FROM PSYCHOLOGY

A person’s experiential life is a stream of conscious experiences. The person’s experienced ‘self’ is *part* of this stream of consciousness: it is not an extra thing that lies outside what the person is conscious of. In James’s words (1890, p. 401) ‘*thought is itself the thinker*, and psychology need not look beyond’. The experiential ‘self’ is a slowly changing ‘fringe’ part of the stream of consciousness. This part of the stream of consciousness provides an overall background cause for the central focus of attention.

The physical brain, evolving mechanically in accordance with the local deterministic process 2, can do most of the necessary work of the brain. It can do the job of creating, on the basis of its interpretation of the clues provided by the senses, a suitable response, which will be controlled by a certain pattern of neural or brain activity that acts as a *template for action*. However, owing to its quantum character, the brain necessarily generates an amorphous mass of overlapping and conflicting templates for action. Process 1 acts to extract from this jumbled mass of possibilities some particular template for action. This template is a feature of the ‘Yes’ states *PSP* that specifies the form of the process 1 event. But the quantum rules do not assert that this ‘Yes’ part of the prior state *S* necessarily comes into being. They assert, instead, that if this process 1 action is triggered (for example, by some sort of ‘consent’) then this ‘Yes’ component *PSP* will come into being with probability $\text{Tr } PSP / \text{Tr } S$, and that the ‘No’ state will occur if the ‘Yes’ state does not occur, where the symbol Tr represents a quantum mechanical summation over all possibilities.

If the rate at which these ‘consents’ occur is assumed to be *increasable by conscious mental effort*, then the causal efficacy of ‘will’ can be understood. Conscious effort can, by activation of the QZE, override strong mechanical forces arising from process 2 and *cause* the template for action to be held in place longer than it would be if the rapid sequence of process 1 events were not occurring. This sustained existence of the template for action can increase the probability that the intended action will occur.

Does this quantum-physics-based concept of the origin of the causal efficacy of ‘will’ accord with the findings of psychology?

Consider some passages from *Psychology: the briefer course*, written by William James. In the final section of the chapter on attention, James (1892, p. 227) writes:

I have spoken as if our attention were wholly determined by neural conditions. I believe that the array of *things* we can attend to is so determined. No

object can *catch* our attention except by the neural machinery. But the *amount* of the attention which an object receives after it has caught our attention is another question. It often takes effort to keep the mind upon it. We feel that we can make more or less of the effort as we choose. If this feeling be not deceptive, if our effort be a spiritual force, and an indeterminate one, then of course it contributes coequally with the cerebral conditions to the result. Though it *introduce* no new idea, it will deepen and prolong the stay in consciousness of innumerable ideas which else would fade more quickly away.

In the chapter on will, in the section entitled ‘Volitional effort is effort of attention’, James (1892, p. 417) writes:

Thus we find that we reach the heart of our inquiry into volition when we ask by what process is it that the thought of any given action comes to prevail stably in the mind.

And, later:

The essential achievement of the will, in short, when it is most ‘voluntary,’ is to attend to a difficult object and hold it fast before the mind.... Effort of attention is thus the essential phenomenon of will.

Still later, James says:

Consent to the idea’s undivided presence, this is effort’s sole achievement.... Everywhere, then, the function of effort is the same: to keep affirming and adopting the thought which, if left to itself, would slip away.

This description of the effect of will on the course of mental–cerebral processes is remarkably in line with *what had been proposed independently from purely theoretical considerations of the quantum physics of this process*. The connections specified by James are explained on the basis of the same dynamic principles that had been introduced by physicists to explain atomic phenomena. Thus the whole range of science, from atomic physics to mind–brain dynamics, has the possibility of being brought together into a single rationally coherent theory of an evolving cosmos that is constituted not of *matter* but of *actions* by agents. In this conceptualization of nature, agents could naturally evolve in accordance with the principles of natural selection, owing to the fact that their efforts have physical consequences. The outline of a possible rationally coherent understanding of the connection between mind and matter begins to emerge.

In the quantum theory of mind/consciousness–brain being described here, there are altogether three processes. First, there is the purely mechanical process called process 2. As discussed at length in the book, *Mind, matter, and quantum mechanics* (Stapp 1993/2003, p. 150), this process, as it applies to the brain, involves important dynamic units that are represented by complex patterns of brain activity that are ‘facilitated’ (i.e. strengthened) by use and are such that each unit tends to be activated as a whole by the activation of several of its parts. The activation of various of these complex patterns by cross referencing—that is, by activation of several of its parts—coupled to feedback loops that strengthen or weaken the activities of

appropriate processing centres, appears to account for the essential features of the mechanical part of the dynamics in a way that is not significantly different from what a classic model can support, except for the existence of a host of parallel possibilities that according to the classic concepts, cannot exist simultaneously.

The second process, von Neumann's process 1, is needed in order to pick out from a chaotic continuum of overlapping parallel possibilities some particular discrete possibility and its complement. (The complement can be further divided, but the essential action is present in the choice of one particular 'Yes' state $PS(t)P$ from the morass of possibilities in which it is imbedded.) The third process is nature's choice between 'Yes' and 'No'. Nature's choice conforms to a statistical rule, but the agent's choice is, within contemporary quantum theory, a 'free choice' that can be and is consistently treated as an input variable of the empirical protocol.

Process 1 has itself two modes. The first is passive, and can produce temporally isolated events. The second is active and involves mental effort.

Active process 1 intervention has, according to the quantum model described here, a distinctive form. It consists of a sequence of intentional purposeful actions, the rapidity of which can be increased with effort. Such an increase in attention density, defined as an increase in the number of observations per unit time, can bring into play the QZE, which tends to hold in place both those aspects of the state of the brain that are fixed by the sequence of intentional actions and also the felt intentional focus of these actions. Attention density is not controlled by any physical rule of orthodox contemporary quantum theory, but is taken both in orthodox theory and in our model, to be subject to subjective volitional control. This application in this way of the basic principles of physics to neuroscience constitutes our model of the mind–brain connection.

(a) *Support from psychology of attention*

A huge amount of empirical work on attention has been done since the nineteenth century writings of William James. Much of it is summarized and analysed in Harold Pashler's (1998) book *The psychology of attention*. Pashler organizes his discussion by separating perceptual processing from post-perceptual processing. The former type covers processing that, first of all, identifies such basic physical properties of stimuli as location, colour, loudness and pitch and, secondly, identifies stimuli in terms of categories of meaning. The post-perceptual process covers the tasks of producing motor actions and cognitive action beyond mere categorical identification. Pashler emphasizes that the empirical 'findings of attention studies... argue for a distinction between perceptual attentional limitations and more central limitations involved in thought and the planning of action' (p. 33). The existence of these two different processes with different characteristics is a principal theme of Pashler's book (e.g. pp. 33, 263, 293, 317, 404).

A striking difference that emerges from the analysis of the many sophisticated experiments is that the perceptual processes proceed essentially in parallel,

whereas the post-perceptual processes of planning and executing actions form a single queue. This is in line with the distinction between 'passive' and 'active' processes. The former are essentially a passive stream of essentially isolated process 1 events, whereas the 'active' processes involve effort-induced rapid sequences of process 1 events that can saturate a given capacity. This idea of a limited capacity for serial processing of effort-based inputs is the main conclusion of Pashler's book. It is in accord with the quantum-based model, supplemented by the condition that there is a limit to how many effortful process 1 events per second a person can produce during a particular stage of their development.

Examination of Pashler's book shows that this quantum model accommodates naturally all of the complex structural features of the empirical data that he describes. Of key importance is his chapter 6, in which he emphasizes a specific finding: strong empirical evidence for what he calls a central processing bottleneck associated with the attentive selection of a motor action. This kind of bottleneck is what the quantum-physics-based theory predicts: the bottleneck is precisely the single linear sequence of mind–brain quantum events that von Neumann quantum theory describes.

Pashler describes four empirical signatures for this kind of bottleneck and describes the experimental confirmation of each of them (p. 279). Much of part II of Pashler's book is a massing of evidence that supports the existence of a central process of this general kind.

The queuing effect is illustrated in a nineteenth century result described by Pashler: mental exertion reduces the amount of physical force that a person can apply. He notes that 'This puzzling phenomenon remains unexplained' (p. 387). However, it is an automatic consequence of the physics-based theory: creating physical force by muscle contraction requires an effort that opposes the physical tendencies generated by the Schrödinger equation (process 2). This opposing tendency is produced by the QZE and is roughly proportional to the number of bits per second of central processing capacity that is devoted to the task. So, if part of this processing capacity is directed to another task, then the applied force will diminish.

The important point here is that there is in principle, in the quantum model, an essential dynamic difference between the unconscious processing done by the Schrödinger evolution, which generates by a local process an expanding collection of classically conceivable experiential possibilities and the process associated with the sequence of conscious events that constitute the wilful selection of action. The former are not limited by the queuing effect, because process 2 simply develops all of the possibilities in parallel. Nor is the stream of essentially isolated passive process 1 events thus limited. It is the closely packed active process 1 events that can, in the von Neumann formulation, be limited by the queuing effect.

The very numerous experiments cited by Pashler all seem to be in line with the quantum approach. It is important to note that this bottleneck is not automatic within classic physics. A classic model could easily produce, simultaneously, two responses in different

modalities, say vocal and manual, to two different stimuli arriving via two different modalities, say auditory and tactile: the two processes could proceed via dynamically independent routes. Pashler notes that the bottleneck is undiminished in split-brain patients performing two tasks that, at the level of input and output, seem to be confined to different hemispheres (p. 308). This could be accounted for by the necessarily non-local character of the projection operator P .

An interesting experiment mentioned by Pashler involves the simultaneous tasks of doing an IQ test and giving a foot response to a rapidly presented sequence of tones of either 2000 or 250 Hz. The subject's mental age, as measured by the IQ test, was reduced from adult to 8 years (p. 299). This result supports the prediction of quantum theory that the bottleneck pertains to both 'intelligent' behaviour, which requires complex effortful processing, and the simple wilful selection of a motor response.

Pashler also notes that 'Recent results strengthen the case for central interference even further, concluding that memory retrieval is subject to the same discrete processing bottleneck that prevents simultaneous response selection in two speeded choice tasks' (p. 348).

In the section on 'mental effort', Pashler reports that 'incentives to perform especially well lead subjects to improve both speed and accuracy', and that the motivation had 'greater effects on the more cognitively complex activity' (p. 383). This is what would be expected if incentives lead to effort that produces increased rapidity of the events, each of which injects into the physical process, through quantum selection and reduction, bits of control information that reflect mental evaluation. Pashler notes 'Increasing the rate at which events occur in experimenter-paced tasks often increases effort ratings without affecting performance. Increasing incentives often raises workload ratings and performance at the same time' (p. 385). All of these empirical connections are in line with the general principle that effort increases attention density, with an attendant increase in the rate of directed conscious events, each of which inputs a mental evaluation and a selection or focusing of a course of action.

Additional supporting evidence comes from the studies of the stabilization or storage of information in short-term memory (STM). According to the physics-based theory, the passive aspect of conscious process merely actualizes an event that occurs in accordance with some brain-controlled rule and this rule-selected process then develops automatically, with perhaps some occasional monitoring. Thus, the theory would predict that the process of stabilization or storage in STM of a certain sequence of stimuli should be able to persist undiminished while the central processor is engaged in another task. This is what the data indicate. Pashler remarks that 'These conclusions contradict the remarkably widespread assumption that short-term memory capacity can be equated with, or used as a measure of, central resources' (p. 341). In the theory outlined here, STM is stored in patterns of brain activity, whereas consciously directed actions are associated with the active selection of a sub-ensemble of quasi-classic states. This distinction seems to account for the large amount of detailed data that

bear on this question of the relationship of the stabilization or storage of information in STM to the types of task that require wilfully directed actions (pp. 337–341). In marked contrast to STM function, storage or retrieval of information from long-term memory (LTM) is a task that requires actions of just this sort (pp. 347–350).

Deliberate storage in, or retrieval from, LTM requires wilfully directed action and hence conscious effort. These processes should, according to the theory, use part of the limited processing capacity and hence be detrimentally affected by a competing task that makes sufficient concurrent demands on the central resources. On the other hand, 'perceptual' processing that involves conceptual categorization and identification without wilful conscious selection should not be interfered with by tasks that do consume central processing capacity. These expectations are what the evidence appears to confirm: 'the entirety of... front-end processing are modality specific and operate independent of the sort of single-channel central processing that limits retrieval and the control of action. This includes not only perceptual analysis but also storage in STM and whatever processing may feed back to change the allocation of perceptual attention itself' (p. 353).

Pashler speculates on the possibility of a neurophysiological explanation of the facts he describes, but notes that the parallel versus serial distinction between the two mechanisms leads, in the classic neurophysiological approach, to the questions of what makes these two mechanisms so different, and what the connection between them is (pp. 354–356, 386–387).

After considering various possible mechanisms that could cause the central bottleneck, Pashler concludes that 'the question of why this should be the case is quite puzzling' (pp. 307–308). Thus, the fact that this bottleneck and its basic properties seem to follow automatically from the same laws that explain the complex empirical evidence in the fields of classic and quantum physics means that the theory being presented here has significant explanatory power for the experimental data of cognitive psychology. Further, it coherently explains aspects of the data that have heretofore not been adequately addressed by currently applicable theoretical perspectives.

These features of the phenomena may be claimed by some to be potentially explainable within a classical-physics-based model. But the possibility of such an explanation is profoundly undermined by the absence from classic physics of the notion of conscious choice and effort. These consciousness-connected features, so critical to a coherent explanation of the psychology of human attention, however, already exist and are specified features of the causal structure of fundamental contemporary physical theory.

7. APPLICATION TO NEUROPSYCHOLOGY

The quantum model is better suited to the analysis of neuropsychological data than models based on the classic approximation. For, just as in the treatment of atomic systems, the quantum approach brings the *phenomenologically described data* directly into the dynamics in place of microscopic variables that are, in

principle, unknowable. Quantum theory injects directly into the causal structure the phenomenal descriptions that we human beings use in order to communicate to our colleagues the empirical facts. It thereby specifies a useful and testable causal structure, while evading the restrictive classic demand that the causal process be ‘bottom up’, i.e. expressible in terms of local mechanical interactions between tiny mindless entities. The Heisenberg uncertainty principle renders that ideal *unachievable in principle* and the banishment of that microlocal ‘bottom up’ determinism opens the door to the quantum alternative of injecting the phenomenologically described realities directly into the causal structure in the way that is both allowed and described by contemporary physical theory.

Quantum physics works better in neuropsychology than its classic approximation precisely because it inserts knowable choices made by human agents into the dynamics in place of unknowable-in-principle microscopic variables. To illustrate this point we apply the quantum approach to the experiment of Ochsner *et al.* (2002).

Reduced to its essence, this experiment consists first of a training phase in which the subject is taught how to distinguish, and respond differently to, two instructions given while viewing emotionally disturbing visual images: ‘attend’ (meaning passively ‘be aware of, but not try to alter, any feelings elicited by’) or ‘reappraise’ (meaning actively ‘reinterpret the content so that it no longer elicits a negative response’). Second, the subjects perform these mental actions during brain data acquisition. The visual stimuli, when passively attended to, activate limbic brain areas, and when actively reappraised, activate prefrontal cerebral regions.

From the classic materialist point of view this is essentially a conditioning experiment where, however, the ‘conditioning’ is achieved through linguistic access to cognitive faculties. But how do the cognitive realities involving ‘knowing’, ‘understanding’ and ‘feeling’ arise out of motions of the miniature planet-like objects of classic physics, which have no trace of any experiential quality? How do the vibrations in the air that carry the instructions get converted into feelings of understanding? And how do these feelings of understanding get converted to conscious effort, the presence or absence of which determine whether the limbic or frontal regions of the brain will be activated?

Within the framework of classic physics these connections between feelings and brain activities remain huge mysteries. The materialist claim (Karl Popper called this historicist prophecy ‘*promissory materialism*’) is that *someday* these connections will be understood. But the question is whether these connections should reasonably be expected to be understood in terms of a physical theory that is known to be false, and to be false in ways that are absolutely and fundamentally germane to the issue. The classic concept demands that the choices made by human agents about how they will act be determined by microscopic variables that according to quantum theory are indeterminate in principle. The reductionist demand that the course of human experience be determined by local mechanical processes is the very

thing that is most conclusively ruled out by the structure of natural phenomena specified by contemporary physical theory. To expect the mind–brain connection to be understood within a framework of ideas so contrary to the principles of physics is scientifically unsupportable and unreasonable.

There are important similarities and also important differences between the classic and quantum explanations of the experiments of Ochsner *et al.* (2002). In both approaches the atomic constituents of the brain can be conceived to be collected into nerves and other biological structures and into fluxes of ions and electrons, which can all be described reasonably well in essentially classic terms. In the classic approach the dynamics must in principle be describable in terms of the local deterministic classic laws that, according to those principles, are supposed to govern the motions of the atomic-sized entities.

The quantum approach is fundamentally different. In the first place the idea that all causation is *fundamentally mechanical* is dropped as being prejudicial and unsupported either by direct evidence or by contemporary physical theory. The quantum model of the human person is essentially dualistic, with one of the two components being described in psychological language and the other being described in physical terms. The empirical/phenomenal evidence coming from subjective reports is treated as data pertaining to the psychologically described component of the person, whereas the data from objective observations, or from measurements made *upon that person*, are treated as conditions on the physically described component of the person. The apparent causal connection *manifested in the experiments* between these two components of the agent is then explained by the causal connections between these components *that are specified by the quantum laws*.

The quantum laws, insofar as they pertain to empirical data, are organized around *events* that increase the amount of information lodged in the psychologically described component of the theoretical structure. The *effects* of these psychologically identified events upon the physical state of the associated brain are specified by process 1 (followed by ‘Nature’s statistical choice’ of which the discrete options specified by process 1 will be experienced). When no effort is applied, the temporal development of the body/brain will be approximately in accord with the principles of classic *statistical* mechanics, for reasons described earlier in connection with the strong *decoherence* effects. But important departures from the classic statistical predictions can be caused by conscious effort. This effort can cause to be held in place for an extended period, a pattern of neural activity that constitutes a *template for action*. This delay can tend to cause the specified action to occur. In the experiments of Ochsner the effort of the subject to ‘reappraise’ *causes* the ‘reappraise’ template to be held in place and the holding in place of this template *causes* the suppression of the limbic response. These causal effects are, by the QZE, mathematical consequences of the quantum rules. Thus the ‘subjective’ and ‘objective’ aspects of the data are tied together by quantum rules that *directly specify the causal effects upon the subject’s brain of the choices made by the subject, without needing to specify how*

these choices came about. The form of the quantum laws accommodates a natural dynamic breakpoint between the *cause* of wilful action, which is not specified by the theory, and its *effects*, which are specified by the theory.

Quantum theory was designed to deal with cases in which the conscious action of an agent—to perform some particular probing action—enters into the dynamics in an essential way. Within the context of the experiment by Ochsner *et al.* (2002), quantum theory provides, via the process 1 mechanism, an explicit means whereby the successful effort to ‘rethink feelings’ actually causes—by catching and actively holding in place—the prefrontal activations critical to the experimentally observed deactivation of the amygdala and orbitofrontal cortex. The resulting *intention-induced modulation* of limbic mechanisms that putatively generate the frightening aversive feelings associated with passively attending to the target stimuli is the key factor necessary for the achievement of the emotional self-regulation seen in the active cognitive reappraisal condition. Thus, within the quantum framework, the causal relationship between the mental work of mindfully reappraising and the observed brain changes presumed to be necessary for emotional self-regulation is *dynamically* accounted for. Furthermore, and crucially, it is accounted for in ways that fully allow for communicating to others the means used by living human experimental subjects to attain the desired outcome. The classic materialist approach to these data, as detailed earlier in this article, by no means allows for such effective communication. Analogous quantum mechanical reasoning can of course be used *mutatis mutandis* to explain the data of Beauregard *et al.* (2001) and related studies of self-directed neuroplasticity (see Schwartz & Begley 2002).

8. CONCLUSIONS

Materialist ontology draws no support from contemporary physics and is in fact contradicted by it. The notion that all physical behaviour is explainable in principle solely in terms of a local mechanical process is a holdover from physical theories of an earlier era. It was rejected by the founders of quantum mechanics, who introduced, crucially into the basic dynamical equations, choices that are not determined by local mechanical processes, but are rather attributed to human agents. These orthodox quantum equations, applied to human brains in the way suggested by John von Neumann, provide for a causal account of recent neuropsychological data. In this account brain behaviour that appears to be caused by mental effort is actually caused by mental effort: the causal efficacy of mental effort is no illusion. Our wilful choices enter neither as redundant nor epiphenomenal effects, but rather as fundamental dynamical elements that have the causal efficacy that the objective data appear to assign to them.

A shift to this pragmatic approach that incorporates agent-based choices as primary empirical input variables may be as important to progress in neuroscience and psychology as it was to progress in atomic physics.

The work of the second-named author (H.P.S.) was supported in part by the Director, Office of Science, Office

of High Energy and Nuclear Physics, Division of High Energy Physics, of the US Department of Energy under Contract DE-AC03-76SF00098. The work of the third-named author (M.B.) was supported in part by a scholarship from the Fonds de la Recherche en Santé du Québec (FRSQ). We thank Joseph O’Neill for his insightful comments about preliminary versions of our manuscript.

APPENDIX A. OTHER INTERPRETATIONS

This work is based on the Copenhagen Interpretation of quantum theory, and von Neumann’s extension of it. The Copenhagen Interpretation is what is, in essence, both taught in standard quantum physics courses and used in actual practice. This interpretation brings the human agent into the dynamics at a crucial point, namely to resolve the ‘basis’ problem, i.e. to pose some particular physical question. This entry of the agent’s ‘free’ choice is the basis of the present work.

Many of the physicists who are most acutely interested in the logical and physical foundations of quantum theory reject the Copenhagen Interpretation, which is basically epistemological and have sought to invent alternative formulations that can be regarded as descriptions of what is really going on in nature. The three most-discussed alternatives are the ones associated with the names of Roger Penrose, Hugh Everett and David Bohm. In order to provide a broader conceptual foundation for understanding and assessing the Copenhagen–von Neumann approach used here we shall compare it with those other three.

All three of the alternative approaches accept von Neumann’s move of treating the entire physical world quantum mechanically. In particular, the bodies and brains of the agents are treated as conglomerations of such things as quantum mechanically described electrons, ions and photons.

Penrose (1994) accepts the need for conscious-related process 1 events, and wants to explain *when* they occur. He proposes an explanation that is tied to another quantum mystery, that of quantum gravity.

Suppose the quantum state of a brain develops two components corresponding to the ‘Yes’ and ‘No’ answers to some query. Penrose proposes a rule, based on the gravitational interaction between these two parts, that specifies approximately how long before a collapse will occur to one branch or to the other. In this way the question of *when* the answer ‘Yes’ or ‘No’ occurs is given a ‘physical’ explanation. Penrose and his collaborator Hameroff (1996) calculate estimates of this typical time-interval on the basis of some detailed assumptions about the brain. The result is a time of the order of one-tenth of a second. They argue that the rough agreement of this number with time-intervals normally associated with consciousness lends strong support to their theory.

The Penrose–Hameroff model requires that the quantum state of the brain has a property called *macroscopic quantum coherence*, which needs to be maintained for around a tenth of a second. But, according to calculations made by Max Tegmark (2000), this property ought not to hold for more than about 10^{-13} s. Hameroff and co-workers (Hagen *et al.* 2002) have advanced reasons why this number should actually be of the order of a tenth of a second. But 12

orders of magnitude is a very big difference to explain away and serious doubts remain about whether the Penrose–Hameroff theory is technically viable.

If *all* aspects of the collapse process were similarly determined in an essentially mechanical way, then there would be in quantum mechanics, as in classic physics, nothing for consciousness to do. But Penrose (1994) argues that the effects of consciousness cannot be purely algorithmic: it cannot be governed by a finite set of rules and operations. His argument is based on the famous incompleteness theorem of Gödel. However, the logical validity of his argument has been vigorously challenged by many experts, including Hillary Putnam (1994), and Penrose's conclusion cannot be deemed absolutely secure. Also, it is peculiar that the question of *when* the event occurs should be essentially algorithmic, while the process itself is non-algorithmic.

However, Penrose's overall aims are similar to those of the approach made in this paper, namely to recognize that the process-1-related features of quantum mechanics are dynamically very different from the local mechanistic dynamics of classic mechanics, or of its quantum analogue, process 2. This differing character of process 1, which is closely connected to conscious awareness, seems, on its face, to be signalling the entry of an essentially non-mechanical consciousness-related element into brain dynamics.

Everett (1957) proposed another way to deal with the problem of how the quantum formulae are tied to our conscious experiences. It is called the many-worlds or many-minds approach. The basic idea is that nature makes no choices between the 'Yes' and 'No' possibilities: both options actually do occur. But, owing to certain features of quantum mechanics, the two streams of consciousness in which these two alternative answers appear are dynamically independent: neither one has any effect on the other. Hence the two incompatible streams exist in parallel *epistemological* worlds, although in the one single ontological or physical quantum world.

This many-minds approach is plausible within the framework provided by quantum mathematics. It evades the need for any real choices between the 'Yes' and 'No' answers to the question posed by the process 1 action. However, von Neumann never even mentions any real choice between 'Yes' and 'No' and the founders of quantum theory likewise focus attention on the crucial *choice of which question shall be posed*. It is *this* choice, which is in the hands of the agent, that the present paper has focused upon. The subsequent choice between 'Yes' and 'No', is normally deemed to be made by nature. But it is enough that the latter choice merely *seems to be made* in accordance with the quantum probably rules. The real problem with the many-minds approach is that its proponents have not yet adequately explained how one can evade the process 1 choices. This difficulty is discussed in detail in Stapp (2002).

David Bohm's pilot-wave model (Bohm 1952) seems at first to be another way of evading the problem of how to tie the formulae of quantum mechanics to human experiences. Yet in David Bohm's book with Basil Hiley (Bohm & Hiley 1993) the last two chapters

go far beyond the reasonably well-defined pilot-wave model and attempt to deal with the problem dealt with in the works of Stapp (1990) and of Gell-Mann & Hartle (1989). This leads Bohm into a discussion of his concept of the implicate order, which is far less mathematically well-defined than his pilot-wave model.

Bohm saw a need to deal with consciousness and wrote a detailed paper on it (Bohm 1986, 1990). His proposals go far beyond the simple well-defined pilot-wave model. It involves an infinite tower of pilot waves, each controlling the level below. The engaging simplicity of the original pilot-wave model is lost in this infinite tower.

The sum of all this is that the structure of quantum theory indicates the need for a non-mechanistic consciousness-related process, but that the approaches to quantum theory that go beyond the pragmatic Copenhagen–von Neumann approach have serious problems that have yet to be resolved. We, in this paper, have chosen to stay on the safer ground of orthodox pragmatic quantum theory and to explore what can be said within that framework.

However, in this addendum we will now stray very briefly from our strict adherence to the pragmatic stance, in order to get a glimpse into what seems to us to be the pathway beyond contemporary pragmatic science that is pointed to by the structure of contemporary physics.

The core message of quantum theory appears to be that the basic realities are 'knowables' not 'be-ables': they are things that can be known, not realities that exist yet cannot be known. This conclusion can be strongly defended by a detailed analysis of quantum theory, but this is not place to do so. However, given this premise, the programme of passing from the anthropocentric set of pragmatic rules to a conception of the greater reality in which our streams of human consciousness are imbedded takes on a different complexion. If nature is constructed of knowables, then the acts of knowing with which we are familiar should be special cases of a pervasive set of similar acts: the world should somehow be constructed of such acts, and of a substrate that is suited to be acted upon by such acts, but that supports, as a matter of principle, only what can become known by other acts. Acts of knowing become, then, the primitives of nature, along with the substrate upon which they act. Conscious acts of probing must also be encompassed.

This way of understanding the meaning of quantum theory opens the door, in principle, to the formulation of rules that allow the choices of which probing actions are taken to depend not exclusively on the current condition of the physically/mathematically described substrate—that is, the current state of the brain—but also on prior acts of knowing. Thus it could well be, as James's remarks suggest, that a mechanical rule determines which thought is *initially caught*, but that felt properties of the consequent act of knowing can influence the rapidity of follow-up repetitions of the probing action.

It is not our intention to propose in this appendix some testable proposal along these lines. We merely note that quantum theory seems naturally to point to this route for an understanding of the reality that lies behind our human-knowledge-based science.

REFERENCES

- Beauregard, M., Lévesque, J. & Bourgouin, P. 2001 Neural correlates of the conscious self-regulation of emotion. *J. Neurosci.* **21**, 1–6.
- Bohm, D. 1952 A suggested interpretation of quantum theory in terms of hidden variables. *Phys. Rev.* **85**, 166–179.
- Bohm, D. J. 1986 A new theory of the relationship of mind to matter. *J. Am. Soc. Psychic. Res.* **80**, 113–135.
- Bohm, D. J. 1990 A new theory of the relationship of mind to matter. *Phil. Psychol.* **3**, 271–286.
- Bohm, D. & Hiley, D. J. 1993 *The undivided universe*. London: Routledge.
- Bohr, N. 1958 *Atomic physics and human knowledge*. New York: Wiley.
- Bohr, N. 1963 *Essays 1958–1962 on atomic physics and human knowledge*. New York: Wiley.
- Cataldi, M., Perez-Reyes, E. & Tsien, R. W. 2002 Difference in apparent pore sizes of low and high voltage-activated Ca^{2+} channels. *J. Biol. Chem.* **277**, 45 969–45 976.
- Everett, H. 1957 Relative state formulation of quantum mechanics. *Rev. Mod. Phys.* **29**, 454–462.
- Gell-Mann, M. & Hartle, J. B. 1989 Quantum mechanics in the light of quantum cosmology. In *Proceedings of Third International Symposium Foundation of Quantum Mechanics* (ed. S. Kobayashi). pp. 321–343. Tokyo: Physical Society of Japan.
- Hagen, S., Hameroff, S. & Tuszynski, J. 2002 Quantum computation in brain microtubules: decoherence and biological feasibility. *Phys. Rev.* **E65**, 061901–1–061901–11.
- Hameroff, S. & Penrose, R. 1996 Orchestrated reduction of quantum coherence in brain microtubules: a model for consciousness. *J. Conscious. Stud.* **3**, 36–53.
- Heisenberg, W. 1958 The representation of nature in contemporary physics. *Daedalus* **87**, 95–108.
- James, W. 1890 *The principles of psychology*, vol. I. New York: Dover.
- James, W. 1892 *Psychology: the briefer course* William James: writings 1879–1899. New York: Library of America (1992).
- Lévesque, J., Joannette, Y., Paquette, V., Mensour, B., Beaudoin, G., Leroux, J.-M., Bourgouin, P. & Beauregard, M. 2003 Neural circuitry underlying voluntary self-regulation of sadness. *Biol. Psychiatry* **53**, 502–510.
- Misra, B. & Sudarshan, E. C. G. 1977 The Zeno's paradox in quantum theory. *J. Math. Phys.* **18**, 756–763.
- Musso, M., Weiller, C., Kiebel, S., Muller, S. P., Bulau, P. & Rijntjes, M. 1999 Training-induced brain plasticity in aphasia. *Brain* **122**, 1781–1790.
- Newton, I. 1964 *Principia mathematica*. In *Newton's principia* (ed. F. Cajori). p. 634. Berkeley: University of California Press.
- Nyanaponika, T. 1973 *The heart of Buddhist meditation*. York Beach, ME: Samuel Weiser.
- Nyanaponika, T. 2000 *The vision of Dhamma: Buddhist writings of Nyanaponika Thera*. Seattle, WA: BPS Pariyatti Editions.
- Ochsner, K. N., Bunge, S. A., Gross, J. J. & Gabrieli, J. D. E. 2002 Rethinking feelings: an fMRI study of the cognitive regulation of emotion. *J. Cogn. Neurosci.* **14**, 1215–1229.
- Paquette, V., Lévesque, J., Mensour, B., Leroux, J.-M., Beaudoin, G., Bourgouin, P. & Beauregard, M. 2003 “Change the mind and you change the brain”: effects of cognitive-behavioral therapy on the neural correlates of spider phobia. *NeuroImage* **18**, 401–409.
- Pashler, H. 1998 *The psychology of attention*. Cambridge, MA: MIT Press.
- Pauli, W. 1955 *The influence of archetypal ideas on the scientific theories of Kepler. The Interpretation of nature and the psyche*. London: Routledge & Kegan Paul.
- Penrose, R. 1994 *Shadows of the mind*. New York: Oxford.
- Putnam, H. 1994 Book review of *Shadows of the mind*, by Roger Penrose. *New York Times Book Review*, November 20, p. 7. See <http://www.ams.org/bull/pre-1996-data/199507/199507015.pdf>.
- Schwartz, J. M. 1998 Neuroanatomical aspects of cognitive-behavioural therapy response in obsessive-compulsive disorder: an evolving perspective on brain and behavior. *Br. J. Psychiatry* **173**(Suppl. 35), 39–45.
- Schwartz, J. M. & Begley, S. 2002 *The mind and the brain: neuroplasticity and the power of mental force*. New York: Harper Collins.
- Schwartz, J. M., Stoessel, P. W., Baxter, L. R., Jr, Martin, K. M. & Phelps, M. E. 1996 Systematic changes in cerebral glucose metabolic rate after successful behavior modification treatment of obsessive-compulsive disorder. *Arch. Gen. Psychiatry* **53**, 109–113.
- Segal, Z. V., Williams, J., Mark, G. & Teasdale, J. D. 2002 *Mindfulness-based cognitive therapy for depression*. New York: Guilford Press.
- Spence, S. S. & Frith, C. 1999 Towards a functional anatomy of volition. *The volitional brain: towards a neuroscience of free will* (ed. B. Libet, A. Freeman & K. Sutherland), pp. 11–29. Thorverton, UK: Imprint Academic.
- Sperry, R. W. 1992 Turnabout on consciousness: a mentalist view. *J. Mind Behav.* **13**, 259–280.
- Stapp, H. P. 1990 Quantum measurement and the mind–brain connection. In *Symposium on the Foundations of Modern Physics* (ed. P. Lahti & P. Mittelstaedt), pp. 403–424. Singapore: World Scientific.
- Stapp, H. P. 1993/2003 *Mind, matter, and quantum mechanics*. New York: Springer-Verlag.
- Stapp, H. P. 2002 The basis problem in many-worlds theories. *Can. J. Phys.* **80**, 1043–1052.
- Tegmark, M. 2000 Importance of quantum decoherence in brain process. *Phys. Rev.* **E61**, 4194–4206.
- Toga, A. W. & Mazziotta, J. C. 2000 *Brain mapping the systems*. San Diego: Academic Press.
- von Neumann, J. 1955 *Mathematical foundations of quantum theory*. Princeton: Princeton University Press.