QUANTUM APPROACHES TO CONSCIOUSNESS.

us

1. Introduction.

Quantum approaches to consciousness are sometimes said to be motivated simply by the idea that quantum theory is a mystery and consciousness is a mystery, so perhaps the two are related. That opinion betrays a profound misunderstanding of the nature of quantum mechanics, which consists fundamentally of a pragmatic scientific solution to the problem of the connection between mind and matter.

The key philosophical and scientific achievement of the founders of quantum theory was to forge a rationally coherent and practically useful linkage between the two kinds of descriptions 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 to us 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 descriptions by causal laws, and 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.

The magnitude of the difference between the quantum and classical conceptions of the connection between mind and brain can scarcely be exaggerated. All approaches to this problem based on the precepts of classical physics founder first on the problem of the lack of any need within classical mechanics for consciousness to exist at all, and second on the seemingly manifest impossibility of ever actually understanding how the experiential realities that form our streams of consciousness could ever be produced by, or naturally come to be associated with, the motions of the things that classical physics claims the physical world to be made of. The first problem is that, according to precepts of classical physics, the causal properties of the physical world suffice, by themselves, to completely specify all physical properties of the universe, including the activities of our bodies and brains, without ever acknowledging the existence of consciousness: everything would go on just the same if nothing but the physical properties were present. The second problem is that the differences-in-kind between the experiential and physical sorts of stuff is so great that it seems beyond the realm of possibility that a tight rational connection could exist between them. The fact that consciousness does exist thus enforces an awkward departure of science from a purely naturalistic stance: nonphysical features such as conscious thoughts, ideas, and feelings must be added to the physically described ones for no apparent naturalistic or physical reason.

Both of these difficulties are resolved in a rationally coherent and practically useful way by quantum mechanics. On the one hand, a key basic precept of the quantum approach, as it is both practiced and taught, is that choices made by human beings play a key and irreducible role in the dynamics. On the other hand, the great disparity within classical physics between the experiential and physical aspects of nature is resolved in the quantum approach by altering the assumptions about the nature of the physical universe. The physical world, as it appears in the theory, is transformed from a structure based on *substance* or *matter* to one based on *events*, each of which has both experiential aspects and physical aspects: Each such event injects information, or “knowledge”, into an information-bearing mathematically described physical state. An important feature of this radical revamping of the conceptual foundations is that it leaves unchanged, at the practical level, most of classical physics. Apart from making room for, and a need for, efficacious conscious choices, the radical changes introduced at the foundational level by quantum mechanics preserve at the pragmatic level almost all of classical physics.

In the remainder of this introductory section I shall sketch out the transition from the classical-physics conception of reality to von Neumann’s application of the principles of quantum physics to our conscious brains. In succeeding sections I describe the most prominent of the many efforts now being made by physicists to apply von Neumann’s theory to recent developments in neuroscience.

The quantum conception of the connection between the psychologically and physically described components of scientific practice was achieved by abandoning the classical picture of the physical world that had ruled science since the time of Newton, Galileo, and Descartes. The building blocks of science were shifted from descriptions of the behaviors of tiny bits of mindless matter to accounts of *the actions that we take to acquire knowledge* and of the *knowledge that we thereby acquire*. Science was thereby transformed from its seventeenth century form, which effectively excluded our conscious thoughts from any causal role in the mechanical workings of Nature, to its twentieth century form, which focuses on our active engagement with Nature, and on what we can learn by taking appropriate actions.

Twentieth century developments have thus highlighted the fact that *science is a human activity* that involves us not as passive witnesses of a mechanically controlled universe, but as agents that can freely choose to perform causally efficacious actions. The basic laws of nature, as they are now understood, not only fail to determine how we will act, but, moreover, inject our *choices about how to act* directly into the dynamical equations.

This altered role of conscious agents is poetically expressed by Bohr's famous dictum:

> "In the great drama of existence we ourselves are both actors and spectators." (Bohr, 1963, p. 15: 1958, p. 81)

It is more concretely expressed in statements such as:

> "The freedom of experimentation, presupposed in classical 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}

The most important innovation of quantum theory, from a philosophical perspective, is the fact that it is formulated in terms of an *interaction* between the physically described world and conscious agents who are, *within the causal structure defined by the known physical laws*, *free to choose* which aspect of nature they will probe. This crack, or gap, in the mechanistic world view leads to profound changes in our conception of nature and man's place within it.

Another key innovation pertains to the *nature* of the *stuff* of the physically/mathematically described universe. The switch is succinctly summarized in Heisenberg's famous assertion:

> "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, 1958a)

What the quantum mathematics describes is not the locations of tiny bits of matter. What it described by the mathematics is a causal structure imbedded in space-time that carries or contains information or knowledge, but no material substance. This structure is, on certain occasions, abruptly altered by *discrete events* that inject new information into it. But this carrier structure is not purely passive. It has an active quality. It acts as a bearer of "objective tendencies" or "potentia" or "propensities" for new events to occur. (Heisenberg, 1958b, p. 53).

To appreciate this new conception of the connection between psychologically described empirical part and the mathematically described physical part of the new scientific description of physical phenomena one needs to contrast it with what came before.

The Classical-Physics Approach.

Classical physics arose from the theoretical effort of Isaac Newton to account for the findings of Johannes Kepler and Galileo Galilei. Kepler discovered that the planets move in orbits that depend on the location of other physical objects - such as the sun - but not on the manner or the timings of our observations: minute-by-minute viewings have no more influence on a planetary orbit than daily, monthly, or annual observations. The nature and timings of our observational acts have no effect at all on the orbital motions described by Kepler. Galileo observed that certain falling terrestrial objects have similar properties. Newton then discovered that he could explain *simultaneously* the celestial findings of Kepler and the terrestrial findings of Galileo by postulating, in effect, that all objects in our solar system are composed of tiny planet-like particles whose motions are controlled by *laws* that refer to the relative locations of the various particles, and make no reference to any conscious acts of experiencing. These acts are taken to be simply passive witnessings of macroscopic properties of large conglomerations of the tiny individually-invisible particles.

Newton's laws involve instantaneous action at a distance: each particle has an instantaneous effect on the motion of every other particle, no matter how distant. Newton considered this non-local feature of his theory to be unsatisfactory, but proposed no alternative. Eventually, Albert Einstein, building on ideas of James Clerk Maxwell, constructed a *local* classical theory in which all dynamical effects are generated by contact interactions between mathematically described properties localized at space-time points, and in which no effect is transmitted faster than the speed of light.

All classical-physics models of Nature are *deterministic*: the state of any isolated system at any time is completely fixed by the state of that system at any earlier time. The Einstein-Maxwell theory is deterministic in this sense, and also "local", in the just-mentioned sense that all interactions are via contact interactions between neighboring localized mathematically describable properties, and no influence propagates faster than the speed of light.

By the end of the nineteenth century certain difficulties with the general principles of classical physical theory had been uncovered. One such difficulty was with "black-body radiation." If one analyzes the electromagnetic radiation emitted from a tiny hole in a big hollow heated sphere then it is found that the manner in which the emitted energy is distributed over the various frequencies depends on the temperature of the sphere, but not upon the chemical or physical character of the interior surface of the sphere: the spectral distribution depends neither on whether the interior surface is smooth or rough nor on whether it is metallic or ceramic. This universality is predicted by classical theory, but the specific form of the predicted distribution differs greatly from what is empirically observed.

In 1900 Max Planck discovered a universal law of black-body radiation that matches the empirical facts. This new law is incompatible with the basic principles of classical physical theory, and involves a new constant of Nature, which was identified and measured by Planck, and is called “Planck’s Constant.” By now a huge number of empirical effects have been found that depend upon this constant, and that conflict with the predictions of classical physical theory.

During the twentieth century a theory was devised that accounts for all of the successful predictions of classical physical theory, and also for all of the departures of the predictions of classical theory from the empirical facts. This theory is called quantum theory. No confirmed violation of its principles has ever been found.

The Quantum Approach.

The core idea of the quantum approach is the seminal discovery by Werner Heisenberg that the classical model of a physical system can be considered to be an *approximation* to a quantum version of that model. This quantum version is constructed by replacing each numerical quantity of the classical model by an *action*: by an entity that acts on other such entities, and for which the order in which the actions are performed matters. The effect of this replacement is to convert each point-like particle of the classical conceptualization—such as an electron—to a smeared-out cloudlike structure that evolves, almost always, in accordance with a quantum mechanical law of motion called the Schroedinger equation. This law, like its classical analog, is local and deterministic: the evolution in time is controlled by contact interactions between localized parts, and the physical state of any isolated system at any time is completely determined from its physical state at any earlier time by these contact interactions. The cloud-like structure that represents an individual “particle”, such as an electron, or proton, tends, under the control of the Schroedinger equation, to spread out over an ever-growing region of space, whereas according to the ideas of classical physics an electron always stays localized in a very tiny region.

The local deterministic quantum law of motion is, in certain ways, incredibly accurate: it correctly fixes *to one part in a hundred million* the values of some measurable properties that classical physics cannot predict.

However, this local deterministic quantum law of motion does not correlate directly to human experience. For example, if the state of the universe were to have developed from the big bang solely under the control of the local deterministic Schroedinger equation then the location of the *center* of the moon would be represented in the theory by a structure spread out over a large part of the sky, in direct contradiction to normal human experience.

This smeared-out character of the position of (the center-point of) a macroscopic object, is a consequence of the famous Heisenberg Uncertainty Principle, combined with the fact that tiny uncertainties at the microscopic level usually get magnified over the course of time, *by the Schroedinger equation acting alone*, to large uncertainties in macroscopic properties, such as location.

Thus a mathematical equation—the Schroedinger equation—that is a direct mathematical generalization of the laws of motion of classical physical theory, and that yields many predictions of incomparable accuracy, strongly conflicts with many facts of everyday experience (e.g., with the fact that the apparent location of the center of the moon is well defined to within, say 10 degrees, as observed from a location on the surface of the earth). Contradictions of this kind must be eliminated by a satisfactory formulation of quantum theory.

In order to put the accurate predictions of the quantum mathematics into the framework of a rationally coherent and practically useful physical theory the whole concept of what physical science is was transformed from its nineteenth form—as a theory of the properties of a mechanical model of Nature in which we ourselves are mechanical parts—to a theory of the connection between the physically and psychologically described aspects of actual scientific *practice*. In actual practice we are agents that probe nature in ways of our own choosing, in order to acquire knowledge that we can use. I shall now describe in more detail how this pragmatic conception of science works in quantum theory.

"The Observer" and "The Observed System" in Copenhagen Quantum Theory.

The original formulation of quantum theory is called the Copenhagen Interpretation because it was created by the physicists that Niels Bohr had gathered around him in Copenhagen. A central precept of this approach is that, in any particular application of quantum theory, Nature is to be considered divided into two parts, "the observer" and "the observed system." The observer consists of the stream of consciousness of a human agent, together with the brain and body of that person, and also the measuring devices that he or she uses to probe the observed system.

Each observer describes himself and his knowledge in a language that allows him to communicate to colleagues two kinds of information: *How he has acted* in order to prepare himself - his mind, his body, and his devices - to receive recognizable and reportable data; and *What he learns* from the data he thereby acquires. This description is in terms of the conscious experiences of the agent himself. It is a description of his intentional probing actions, and of the experiential feedbacks that he subsequently receives.

In actual scientific practice the experimenters are free to choose which experiments they perform: the empirical procedures are determined by the protocols and aims of the experimenters. This element of freedom is emphasized by Bohr in statements such as:

> "To my mind there is no other alternative than to admit in this field of experience, we are dealing with individual phenomena and that our possibilities of handling the measuring instruments allow us to make a choice between the different complementary types of phenomena that we want to study. (Bohr, 1958, p. 51)

This freedom to choose is achieved in the Copenhagen formulation of quantum theory by placing the empirically/psychologically described observer outside the observed system that is being probed, and then subjecting only the observed system to the rigorously enforced mathematical laws.

The observed system is, according to both classical theory and quantum theory, describable in terms of mathematical properties assigned to points in space-time. However, the detailed forms of the laws that govern the evolution in time of this mathematical structure, and of the rules that specify the connection of this mathematical structure to the empirical facts, are very different in the two theories.

I am endeavoring here to avoid mathematical technicalities. But the essential conceptual difference between the two approaches rests squarely on a certain technical difference. This difference can be illustrated by a simple two-dimensional picture.

The Paradigmatic Example.

Consider an experiment in which an experimenter puts a Geiger counter at some location with the intention of finding out whether or not this device will "fire" during some specified time interval. The experiment is designed to give one of two possible answers: 'Yes', the counter will fire during the specified interval, or 'No', the counter will not fire during this specified interval. This is the paradigmatic quantum measurement process.

This experiment has *two* alternative mutually exclusive possible responses, 'Yes' or 'No.' *Consequently*, the key mathematical connections can be pictured in a *two*-dimensional space, such as the top of your desk.

Consider two distinct points on the top of your desk called *zero* and *p*. The displacement that would move a point placed on *zero* to the point *p* is called a *vector*. Let it be called *V*. Suppose *V* has unit length in some units, say meters. Consider any two other displacements *V1* and *V2* on the desk top that start from zero, have unit length, and are perpendicular to each other. The displacement *V* can be formed in a unique way by making a (positive or negative) displacement along *V1* followed by a (positive or negative) displacement along *V2*. Let the lengths of these two displacements be called *X1* and *X2*, respectively. The theorem of Pythagoras says that *X1* squared plus *X2* squared is one (unity).

Quantum theory is based on the idea that the various experiencable outcomes have "images" in a vector space. The vector V1 mentioned above is the image, or representation, in the vector space of the possible outcome 'Yes,' whereas V2 represents 'No.' I will not try to describe here how this mapping of possible experiencable outcomes into corresponding vectors is achieved. But the basic presumption in quantum theory is that such a mapping exists.

The vector *V* represents the state of the to-be-observed system, which has been prepared at some earlier time, and has been evolving in accordance with the Schroedinger equation. The vector *V1* represents the state that this observed system would be known to be in if the observed outcome of the measurement were 'Yes.' The vector *V2* represents the state that the observed system would be known to be in if the observed result of the measurement were 'No.' Of course, the directions of the two perpendicular vectors *V1* and *V2* depend upon the exact details of the experiment: on exactly where the experimenters have placed the Geiger counter, and on other details controlled by the experimenters.

The outcome of the probing measurement will be either *V1* (Yes) or *V2* (No). The predicted probability for the outcome to be 'Yes' is *X1* squared and the predicted probability for the outcome to be 'No' is *X2* squared. These two probabilities sum to unity, by virtue of the theorem of Pythagoras. The sudden jump of the state from *V* to either *V1* or *V2* is called a "quantum jump." The general theory is express in terms of a many-dimensional generalization of your desktop. This generalization is called a Hilbert space, and every observable state of a physical system is a represented by a "vector" in such a space.

The crucial, though trivial, logical point can now be stated: The two alternative possible outcomes, 'Yes' or 'No' of the chosen-by-the-experimenter experiment are associated with a pair of perpendicular unit-length vectors called "basis vectors". The *orientation* (i.e., directions) of the set of "basis" vectors, *V1* and *V2,* enters into the dynamics as a *free variable* controlled by the experimental conditions, which are specified in practice by choices made by experimenters. The orientation of the set of basis vectors is thus, from a mathematical standpoint, a variable that can be, and is, specified *independently* of the state *V* of the system being probed.

This entry into the dynamics of choices made by the experimenters is not at all surprising. If the experimenters are considered to stand outside, and apart from, the system being observed, as specified by the Copenhagen approach, then it is completely reasonable and natural that the choices made by the experimenters (about how to probe the observed system) should be treated as variables that are independent of the variables that specify the physical state of the system they are probing.

Bohr (1958: 92, p. 100) argued that quantum theory should not be applied to living systems. He also argued that the classical concepts were inadequate for that purpose. So the strict Copenhagen approach is simply to renounce the applicability of *contemporary* physical theories, both classical and quantum, to neurobiology.

Von Neumann's Formulation.

The great mathematician and logician John von Neumann (1955/1932) rigorized and extended quantum theory to the point of being able to corporate the devices, the body, and the brain of the observers into the physically described part of the theory, leaving, in the psychologically described part, only the stream of conscious experiences of the agents. The part of the physically described system being directly acted upon by a psychologically described "observer" is, according to von Neumann's formulation, *the brain of that observer*. (von Neumann, 1955, p. 421). The quantum jump of the state of the brain of an observer to the 'Yes' basis state (vector) then becomes the representation, *in the state of that brain*, of the conscious acquisition of the knowledge associated with that answer 'Yes.' Thus the physical features of the brain state actualized by the quantum jump to the state V1 associated with the answer 'Yes' constitute the *neural correlate* of that person's conscious experience of the feedback 'Yes.' This fixes the essential quantum link between consciousness and neuroscience.

This is the key point! Quantum physics is built around "events" that have both physical and phenomenal aspects. The events are physical because they are represented in the physical/mathematical description by a "quantum jump" to one or another of the basis state vectors defined by the agent/observer's choice of what question to ask. If the resulting event is such that the 'Yes' feedback experience occurs then this event "collapses" the prior physical state to a new physical state compatible with that phenomenal experience. Mind and matter thereby become dynamically linked in a way that is causally tied to the agent's free choice of how he will act. Thus a causal dynamical connection is established between (1) a person's conscious choices of how to act, (2) his consciously experienced increments in knowledge, and (3) the actualizations of the neural correlates of the experienced increments in knowledge.

This conceptualization of the structure of basic physical theory is radically different from what it was in classical physics. Classical physics was based on a guess that very worked well for two centuries, namely the notion that the concepts that provided an "understanding" of our observations of planets and falling apples would continue to work all the way down to the elementary-particle level. That conjecture worked well until science became able to explore what was happening at the elementary-particle or atomic level. Then it was found that that simple "planetary" idea could not be right. So scientists turned to a more sophisticated approach that was based less on simplistic ontological pre-suppositions and more on the empirical realities of actual scientific practice.

This new conceptual structure is not some wild philosophical speculation. It rationally yields—when combined with the statistical rule associated with the theorem of Pythagoras described above—all the pragmatic results of quantum theory, which include, as special cases, all the valid predictions of classical physics!

Von Neumann shifted the boundary between the observer and the observed system, in a series of steps, until the bodies and brains of all observers, and everything else that classical physics would describe as "physical" was included as part of the observed system, and showed that this form of the theory is essentially equivalent, in practice, to the Copenhagen interpretation. But it evades an unnatural limitation imposed by Bohr: it by-passes the ad hoc separation of the dynamically unified physical world into two differently described parts. Von Neumann's final placement of the boundary allows the psychological description to be—as is natural—the description of a stream of conscious experiences that are the experiential sides of a sequence of events whose physical sides actualize the neural correlates of those experiences.

It is important that von Neumann's systematic enlargement of the physical system to include eventually the bodies and brains of the observers *does not disrupt the basic mathematical structure of the theory.* In particular, it does not alter *the critical need to specify the orientation of the set of basis vectors* (e.g., V1 and V2) in order to make the theory work. *The specification of the basis states continues to be undetermined by anything in contemporary physical theory*, *even when the physical description is extended to include the entire physical world, including the bodies and brains of all human observers.*

This leap by von Neumann from the realm of atomic physics to the realm of neuroscience was way ahead of its time. Neuroscience was then in a relatively primitive state compared to what it is today. It had a long way to go before mainstream interest turned to the question of the connection between brains and conscious experiences. But 70 years of brain science has brought the empirical side up to the level where the details of the mind-brain connections are being actively probed, and intricate results are being obtained that can be compared to the predictions of the psycho-physical theory prepared long ago by John von Neumann.

It is evident that a scientific approach to brain dynamics must *in principle* use quantum theory, in order to deal properly with brain processes that depend heavily on chemical and ionic processes. For example, the release of neurotransmitter from a nerve terminal is controlled by the motions of calcium ions, and these ions are small enough so that the deterministic laws of classical physics necessarily fail. Quantum theory must *in principle* be used to describe the ion dynamics. But then the state of the brain is in principle a cloud-like structure that can encompass many conflicting classical possibilities. The generation, within the quantum state of the brain, of conflicting classical possibilities should occur particularly when the low-level essentially mechanical processes cannot come to agreement on the best course of action. A higher order “executive decision” is needed. It is probably important in this connection that, unlike the mechanical evolution generated by the local deterministic Schoedinger equation, the quantum jumps associated with conscious experiences are intrinsically nonlocal: they can access *together* physical features that are located over extended portions of the brain.

Summary.

The essential difference at the basic conceptual level between the quantum and classical approaches to consciousness is that the classical principles make no mention of consciousness. The causal structure is in principle completely “bottom up.” Everything is, in principle, fully determined by what goes on at the microscopic atomic level, and any dependence of microscopic properties upon macroscopic properties, or on consciousness, is, in the end, a round-about consequence of laws expressible exclusively in terms of properties of atomic particles and of the physical fields that they produce. But in quantum theory the local-deterministic (i.e., bottom-up) physical process is *in principle causally incomplete*. It fixes, by itself, neither our actions nor our experiences, nor even any statistical prediction about how we will act or what we will experience. The bottom-up process *alone* is unable to make statistical predictions, because the statistical predictions depend upon the choice of a set of basis vectors, and the bottom-up local-deterministic quantum process does not fix this choice.

This reorganization of the dynamical structure leads to an altered perspective on the entire scientific enterprise. The psychologically described empirical side of scientific practice is elevated from its formerly subservient status - as something that should be *deduced* from, or constructed from, the already-dynamically-complete physical side - to the new status of co-equal dynamical partner. Science becomes the endeavor to describe the *two-way interplay* between the psychologically and physically described aspects of nature, rather than an attempt to deduce the existence and properties of our streams of conscious experiences from a presumed-to-be-dynamically-complete local mechanical model.

Within the von Neumann framework our conscious choices fix the orientations of the basis vectors. These choices can strongly influence our actions. Thus these influences need not be illusions. The theory provides, as we shall see in the section 4, a specific mechanism that allows our conscious “free” choices to significantly influence our physical actions.

Pragmatic Neuroscience.

Von Neumann, in his 1932 book followed the Copenhagen tack of focusing on scientific practice rather than ontological issues. Indeed, it can be argued that science is intrinsically pragmatic rather than ontological. The true nature of things, other than our experiences themselves, can never be found by the methods of science. Thus Von Neumann's formulation of quantum theory provides the foundations of a *pragmatic* neuro-psycho-dynamics that is built on contemporary physical theory, rather than an inappropriate classical physics. All quantum approaches to consciousness build upon this foundation laid by von Neumann, but various physicists have proposed different ways of developing that core structure. We turn now turn to the descriptions of a number of these proposals.

2. The Penrose-Hameroff Approach.

Perhaps the most ambitious attempt to create a quantum theory of consciousness is the one of Roger Penrose and Stuart Hameroff. Their proposal has three parts: The Gödel Part, The Gravity Part, and the Microtubule Part.

The Gödel Part, which is due to Penrose, is an effort to use the famous Gödel Incompleteness Theorem to prove that human beings have intellectual powers that they could not have if they functioned in accordance with the principles of classical physical theory. Proving this would reaffirm a conclusion of the von Neumann formulation of quantum theory, namely that a conscious human being can behave in ways that a classical mechanical model cannot. Penrose's argument, if valid, would yield this same conclusion, but within a framework that relies not on quantum concepts, which are generally unknown to cognitive scientists, but rather on Gödel-type arguments, which are familiar to some of them.

The general idea of Penrose's argument is to note that, due to the mathematically deterministic character of the laws of classical physics, the output at any specified finite time of any computer behaving in accordance with the classical laws should in principle be deducible, to arbitrarily good accuracy, from a finite-step procedure based on a finite set of mutually consistent rules that encompass the laws of arithmetic. But then a human being who can be adequately modeled as a classical computer should be able to know, at any finite time, the truth *only* of those statements that can be deduced from a finite-step computation based on the finite set of rules that govern that computer. Yet Gödel-theorem-type arguments allow real mathematicians to know, given *any* finite set of consistent logical rules that encompass the laws of arithmetic, the truth of mathematical statements that cannot be deduced by any finite-step proof based on those rules. This seems to imply that a real mathematician can know things that no classical physics model of himself could ever know, namely the truth of statements that his classical computer simulation could not establish in a finite time.

Filling in the details of this argument is not an easy task. Penrose spends the better part of five chapters in "The Emperor's New Mind," (Penrose, 1989) and some 200 pages in "Shadows of the Mind" (Penrose, 1994) explaining and defending this thesis. However, the Harvard philosopher Hillary Putnam challenged Penrose's conclusion in a debate appearing in the New York Times Review of Books, (Putnam, 1994) and numerous logicians have since weighed in, all, to my knowledge, challenging the validity of Penrose's argument. Thus the Gödel Part of the Penrose-Hameroff approach cannot now be regarded as having been successfully established.

The Gravity Part of the Penrose-Hameroff approach addresses a key question pertaining to the quantum dynamics: exactly *when* do the sudden "quantum jumps" occur? In von Neumann's theory these jumps should presumable occur when the neural correlates of conscious thoughts become sufficiently well formed. But von Neumann gives no precise rule for when this happens.

The lack of specificity on this issue of precisely "*when"* is a serious liability of the von Neumann theory, insofar as it is construed as a description of the ontological mind-matter reality itself. That difficulty is the basic reason why both the original Copenhagen formulation and von Neumann's extension of it eschew traditional ontological commitments. They hew rather to the pragmatic position that the job of science is to establish useful practical connections between empirical findings and theoretical concepts, rather than advancing shaky speculations about the ultimate Nature of reality. The pragmatic position is that theoretical ideas that optimally provide reliable practical connections between human experiences constitute, themselves, our best *scientific* understanding of "reality." Added ontological superstructures are viewed as not true science, because additions that go beyond optimal theoretical descriptions of connections between human experiences cannot be tested empirically.

Penrose wants to provide an ontology that has "real quantum jumps." Hence he must face the issue: when do these jumps occur. He seeks to solve this problem by linking it to a problem that arises when one attempts to combine quantum theory with Einstein's theory of gravity.

Einstein's theory of gravity, namely General Relativity, is based of the idea that space-time is not a rigid flat structure, as had previously been thought, but rather a *deformable medium*, and that the way it is deformed is connected to the way that matter is distributed within it. This idea was developed within the framework of classical physical theory, and most applications of it are made within a classical-physics idealization. But serious problems arise when the quantum character of "matter" is considered. For, according to orthodox quantum theory, a particle, such as an electron or an ion, has no well defined location: its location is specified by a smeared out “probability cloud.” But if the locations of the material particles are not well defined then, according to General Relativity, neither is the form of the space-time structure in which the particle structures are imbedded.

Penrose conjectures that Nature abhors uncertainty in the structure of space-time, and that when too much ambiguity arises in the space-time structure a quantum jump to some less ambiguous structure will occur. This “principle” allows him to tie quantum jumps to the amount of uncertainty in the structure of space-time.

There is no compelling reason why Nature should be any more perturbed by an uncertainty in the structure of space-time than by an uncertainty in the distribution of matter. However, by adopting the principle that Nature finds intolerable an *excessive ambiguity in* the *structure of space-time* Penrose is able to propose a specific rule about when the quantum jumps occur.

Penrose’s rule depends on the fact that Planck’s constant gives a connection between energy and time: this constant divided by any quantity of energy gives a corresponding interval of time. Thus if an energy associated with a possible quantum jump can be defined then a time interval associated with that potential jump becomes specified.

To identify the pertinent energy consider a simple case in which, say, a small object is represented quantum mechanically by a small cloud that divides into two similar parts, one moving off to the right, the other moving off to the left. Both parts of the cloud are simultaneously present, and each part produces *a different distortion* of the underlying spacetime structure, because matter is distributed differently in the two cases. One can compute the amount of energy that it would take to pull apart, against their gravitational attraction, two copies of the object, if each copy is located at the position specified by one of the two clouds. If one divides Planck’s constant by this “gravitational energy” then a time interval associated with this distortion of space-time into these two disparate structures becomes defined. Penrose proposes that this time interval is the duration of time for which Nature will *endure* this bifurcation of its space-time structure into the two incompatible parts, before jumping to one or the other of these two forms.

This conjectured rule is based on two very general features of Nature: Planck’s universal constant of action and the Newton-Einstein universal law of gravitation. This universality makes the rule attractive, but no reason is given why Nature must comply with this rule.

Does this rule have any empirical support?

An affirmative answer can be provided by linking Penrose’s rule to Hameroff’s belief that consciousness is closely linked to the *microtubular sub-structure of the neurons.*

It was once thought that the interiors of neurons were basically structureless fluids. That conclusion arose from direct microscopic examinations. But it turns out that in those early studies the internal substructure was wiped out by the fixing agent. It is now known that neurons are filled with an intricate structure of *microtubules*.

Each microtubule is a cylindrical structure that can extend over many millimeters. The surface of the cylinder is formed by a spiral chain of tubulin molecules, with each circuit formed by thirteen of these molecules. The tubulin molecule has molecular weight of about 110,000 and it exists in two slightly different configurational forms. Each tubulin molecule has a single special electron that can be in one of two relatively stable locations. The molecule will be in one or the other of the two configurational states according to which of these two locations this special electron is occupying.

Hameroff is an anesthesiologist, and he noted that there is close correspondence between, on the one hand, the measured effects of various anesthetics upon consciousness and, on the other hand, the capacity of these anaesthetics to diminish the ability of the special electron to move from one stable location to the other. This suggests a possible close connection between consciousness and the configurational activity of microtubules.

This putative linkage allows an empirical test of Penrose’s rule to be made.

Suppose, in keeping with the case considered by Penrose, you are in a situation where one of two possible experiences will probably occur. For example, you might be staring at a Necker Cube, or walking in a dark woods when a shadowy form jumps out and you must choose “fight” or “flight,” or perhaps you are checking your ability to freely choose to raise or not raise your arm. Thus one of two alternative possible experiences is likely to occur. Various experiments suggest that it takes about half a second for an experience to arise. Given this time interval, Penrose’s formula specifies a certain corresponding energy. Then Hameroff can compute, on the basis of available information concerning the two configurational states of the tubulin molecule, how many tubulin-molecule configurational shifts are needed to give this energy.

The answer is about 1% of the estimated number of tubulin molecules in the human brain. This result seems reasonable. Its reasonableness is deemed significant, since the computed fraction could have come out to be perhaps billions of times smaller than, or billions of times greater than, 100%. The fact that the computed value is “in the ballpark” supports the idea that consciousness may indeed be closely connected to tubulin configurational activity.

Given this rather radical idea—it was previously thought that the microtubules were merely a construction scaffolding for the building and maintenance of the physical structure of the neurons—many other exotic possibilities arise. The two configurational forms of the tubulin molecule mean that it can hold a “bit” of information, so maybe the microtubular structure forms the substrate of a complex *computer* located within each neuron, thus greatly expanding the computational power of the brain. And maybe each such computer is in fact a “quantum computer.” And maybe these quantum computers are all linked together to form one giant brain-wide quantum computer. And maybe these hollow micro-tubes form wave guides for quantum waves.

These exotic possibilities are exciting and heady ideas. They go far beyond what conservative physicists are ready to accept, and far beyond what the 1% number derived from Penrose’s rule actually supports. What is supported is merely a connection between consciousness and microtubular activity, *without the presence* of the further stringent *coherence conditions* required for the functioning of a quantum computer.

“Coherence” means preservation of the “phase” relationships that allow waves that have traveled via different paths to come back together so that, for example, crest meets crest and trough meets trough to build an enhanced effect. Quantum computation requires an effective isolation of the quantum informational waves from the surrounding environment, because any interaction between these waves and the environment tends to destroy coherence. But the required isolation is difficult to maintain in a warm, wet, noisy brain.

The simplest system that exhibits a behavior that depends strongly on quantum interference effects, and for which the maintenance of *coherence* is essential, is the famous “double-slit experiment.” When photons of a single wave length are allowed to pass, one at a time, through a pair of closely spaced narrow slits, and each photon is later detected by some small detection device that is imbedded in a large array of such devices, one finds that *if the photonic system is not allowed to perceptibly influence any environmental degree of freedom* on its way to the detection device then the pattern of detected events depends on an *interference* between the parts of the beam passing through the two different slits. This pattern is very different from what it is if the photon is allowed to perceptibly disturb, the surrounding environment. Disturbing the environment produces a “decoherence” effect, i.e., a weakening or disappearance of the interference effects.

If a system interacts with its environment, it is difficult to prevent a “perceptible influence” of the system on the environment. But if even *a single one* of the thousands of particles in the environment is displaced by a discernible amount then the coherence is lost, and the quantum interference effect will disappear.

Since the medium in which the putative quantum information waves are moving involves different conformational states of huge tubulin molecules of molecular weight ~110,000, it would seemingly be exceedingly hard to ensure that the passage of these waves will not disturb even one particle of the environment by a discernible amount.

Max Tegmark wrote an influential paper in Physical Review E. (Tegmark, 2000). It mathematically buttressed the intuition of most physicists that the macroscopic coherence required by Penrose-Hameroff---namely that the microtubular conformal states can form the substrate of a quantum computer that extends over a large part of the brain--- could not be realized in a living human brain. Tegmark concluded that the coherence required for macroscopic quantum computation would be lost in a ten trillionth of a second, and hence should play no role in consciousness. This paper was widely heralded. However, Hagan, Hameroff, and Tuszynski (2002) wrote a rejoinder in a later issue of the same journal. They pointed out several departures of Tegmark's assumptions from those of the Penrose-Hameroff model. The associated corrections lengthened the coherence time by 8 or 9 orders of magnitude, thus bringing the situation into a regime where the non-equilibrium conditions in a living brain might become important: energetic biological processes might conceivably intervene in a way that would make up the still-needed factor of ten thousand. However, the details of how this might happen were not supplied. Hence the issue is, I believe, still up in the air, with no detailed explanation available to show how the needed macroscopic quantum coherence could be maintained in a living human brain.

It must be stressed, however, that these exotic "quantum computer" effects are not necessary for the emergence of strong quantum effects within the general framework supplied by the combination of Penrose's rule pertaining to gravity and Hameroff's claim concerning the importance of microtubules. According to von Neumann's general formulation, the state of the brain—or of the microtubular part of the brain—is adequately represented by what physicists call the "reduced density matrix" of that subsystem. This representation depends only on the variables of that subsystem itself (i.e., the brain, or microtubular array) but nevertheless takes adequate account of the interactions of that system with the environment. It keeps track of the quantum coherence or lack thereof. Penrose's rule can be stated directly in terms of the "reduced density matrix," which displays, ever more clearly as the interaction with the environment grows, the two alternative states of the brain—or of the microtubular array—that Nature must choose between. This reduced-density-matrix representation shows that the powerful decoherence effect produced by strong interactions with the environment actually *aids* the implementation of Penrose's rule, which is designed to specify *when* the quantum jump occurs (and perhaps to which states the jump occurs). The capacity of the brain to be or not to be *a quantum computer* is a very different question, involving enormously more stringent conditions. It thus is important, for logical clarity, to separate these two issues of the requirements for *quantum computation* and for *quantum jumps*, even though they happen to be interlocked in the particular scenario described by Penrose and Hameroff,

3. The Bohm Approach.

The Copenhagen and von Neumann formulations of quantum theory are non-deterministic. Both specify that human choices enter into the dynamics, but neither specifies the causal origins of these choices. The question thus arises: what determines these choices?

One possibility is that these choices arise in some yet-to-be-specified way from what we conceive to be the *idealike aspect of reality*. That option was pursued by Penrose, with his suggestion that our thoughts are linked to Plato's world of ideal forms. Another—seemingly different—possibility is that a *physical description* exists that is more detailed than the smeared out cloudlike structures of the orthodox formulations, and that this *more detailed physical description* determines all features left undetermined in the orthodox formulations.

This second approach was developed by David Bohm (1952, 1993). His formulation of quantum theory postulates, in effect, the existence of the old-fashioned world of classical physical theory. This classical-type world is supposed to exist *in addition to the cloudlike wave function of orthodox quantum theory* and is supposed to evolve in a way completely determined by what precedes it in time. Bohm species new laws of motion that are able to reinstate determinism in a way compatible with the predictions of quantum theory, but at the expense of a very explicit abandonment of locality: Bohm's theory entails very strong, and very long-range, instantaneous action-at-a-distance.

One serious failing of Bohm's approach is that it was originally formulated in a non-relativistic context, and it has not yet – after half a century and great effort – been extended to cover the most important domain in physics, namely the realm of quantum electrodynamics. This is the theory that covers the atoms that make up our bodies and brains, along with the tables, chairs, automobiles, and computers that populate our daily lives. This deficiency means that Bohm's theory is, at present, primarily a philosophically interesting curiosity, not a practically useful physical theory.

Also, Bohm's theory, at least in its original form, is not really germane to the issue of consciousness. For Bohm's theory *successfully achieved its aim,* which was precisely to get rid of consciousness: i.e., to eliminate consciousness from the basic dynamical equations, just as classical physics had done.

Bohm recognized, later on, that some understanding of consciousness was needed, but he was led instead, to the notion of an infinite tower of mechanical levels, each controlling the one below, with consciousness somehow tied to the mystery of the infinite limit. (Bohm, 1986, 1990) This infinite-tower idea tends to diminish the great achievement of the original theory, which was to reinstate physical determinism in a simple way.

To appreciate the consequences of adopting a theory with *complete physical determinism compatible with the empirical predictions of quantum theory* it is instructive to examine Bohm's original deterministic model, and see how, within that deterministic framework in which consciousness plays no fundamental causal role, consciousness nevertheless enters at the level of scientific practice.

As explained in the introductory section, actual scientific practice involves setting up experimental conditions that promote consciously conceived objectives. In von Neumann's theory these consciously chosen actions influence the subsequent course of events in the observed system, which, according to von Neumann's version of quantum theory, is primarily the brain of the human participant. A key point is that these choices, made by the experimenter about how he or she will act, are treated in von Neumann's theory, and also by Copenhagen quantum theory, as *input data*, to be fixed by the experimenter. No matter what these choices actually are, or where they come from, or what they actually do, these conscious choices are *treated* in orthodox quantum theory as free, controllable and knowable, input boundary conditions.

In Bohm's theory these choices are not *actually* free: freedom is an illusion. The apparently free choice is, at a deeper dynamical level, completely determined by *physical* conditions, just as it was in classical physics. However, the putative existence of this deeper dynamical underpinning does not upset scientific practice. It does not displace, within science, the orthodox quantum dynamics. The analysis of Heisenberg shows that, even within the context of a deterministic Bohmian mechanics, the human observers can never determine, or *know*, to which of the conceivable, logically possible classical Bohmian worlds their experiences belong. The Heisenberg Uncertainty Principle is a limitation upon human knowledge that is not evaded by Bohm's deterministic dynamics. The most that "experiencers" can ever *know* about the Bohmian classical world of which they are a putative part is represented by a quantum mechanical cloud-like wave function.

This limitation *in human knowledge* is acknowledged by Bohm. Indeed, Bohm's theory leaves actual scientific practice the same as it is in the Copenhagen approach. This *equivalence at the practical level* of Bohm's model to the Copenhagen formulation means that the unavoidable gap in human knowledge mandated by the uncertainty principle is dealt with by returning to Copenhagen quantum theory. The theoretically specified, *but in principle unknowable and uncontrollable information about the supposedly deterministic microscopic realities* are replaced in actual practice by *knowable and controllable realities*, namely our human conscious choices about which actions we will take, and their consciously experienced feedbacks..

The point here is that the details of the Bohmian microstructure can, as a matter of principle, never be known to us, and hence cannot be used to make predictions. What we can and do experience are our efforts to act in intended ways, and the consequences of our conscious choices about how to act, *These experiential feedbacks place conditions on the putative Bohmian microstructure*. These consciously experiences are what we know, and if we try to use what we know to make predictions about future experiences then the deterministic theory constructed by Bohm is such that the best we can do is to use the Copenhagen statistical rules. Thus the extra infra-structure added by Bohm, while it may allow us to think that we have a truer or better understanding of what is really going on, adds no testable predictions that go beyond what the orthodox theory provides. The orthodox quantum approach is thus to regard as philosophically speculative, and beyond the scope of science, any ascription of reality to properties that are *unknowable in principle*, and to build physical theory upon what is *knowable in principle.* This approach may be as sound and useful in neuroscience as it is in atomic physics.

4. The von Neumann/Stapp Approach

John von Neumann converted Copenhagen quantum theory, in a series of steps, into a form in which the entire physical universe, including the brain of each agent, is represented in one basic quantum state, which is called the state of the universe. The state of any subsystem, such as a brain, is formed by averaging (tracing) this basic state over all variables *other* than those that describe the state of that subsystem. The dynamics consists of *three* processes.

Process **1** is the choice on the part of the experimenter about how to act. This choice is sometimes called "The Heisenberg Choice," because Heisenberg strongly emphasized its crucial role in quantum dynamics. At the pragmatic level it is a "free choice," because it is controlled *in practice* by the conscious intentions of the experimenter/participant, and neither the Copenhagen nor von Neumann formulations provide any description of the *causal origins* of this choice, apart from the thoughts, ideas, and feelings of the agent. Each intentional action involves an effort to produce a conceived experiential feedback, which, if it occurs, will be an experiential confirmation of the success of the intended action

Process **2** is the quantum analog of the equations of motion of classical physics. As in classical physics, these equations of motion are local: all interactions are between immediate neighbors. They are also deterministic. They are obtained from the classical equations by a certain *quantization* procedure, and are reduced to the classical equations by taking the *classical approximation* of setting to zero the value of Planck's constant everywhere it appears. Evolution via the quantum Process **2** normally has the effect of *expanding* the microscopic uncertainties beyond what is demanded by the Heisenberg Uncertainty Principle:

the cloud of microscopic possibilities *spreads out*. This growth in the microscopic regime, if unchecked by any other process, spreads into the macroscopic domain and causes even the *centers* of large objects to tend to become diffused over large regions. The disparity between this Process-**2**-generated theoretical indefiniteness of the locations of the centers of large objects and the consciously experienced definiteness of the positions of visible objects is resolved by Process **3**.

Process **3** is sometimes call the "Dirac Choice." Dirac called it a "choice on the part of Nature." It can be regarded as Nature's answer to the question posed by Process **1**. This posed question might be: "Will the detecting device be found to be in the state that signifies "Yes, a detection has occurred" ? Or, "Will the Geiger counter be observed to 'fire' in accordance with the experiential conditions that define a `Yes' response?" Each Process **3** reply must be preceded by a Process **1** question. This is because the Process **2** generates a *continuous infinity* of possible questions that cannot all be answered consistently within the mathematical framework provided by quantum theory. Process **1** specifies a set of distinct allowed possible answers such that the "Pythagoras Rule" for probabilities yields the conclusion that the probabilities for the allowed possible answers sum to unity.

Process **1** brings the conscious choices made by the observer/participant directly into the dynamics. On the other hand, there is a tendency for the effect of the Process-**1** choices (of the questions) on the state of observed system to be washed out, in the long run, by the averaging over the two possible answers, `Yes' and 'No.' However, it has been stressed by Stapp (1999) that if willful effort can control the *rate* at which a sequence of similar Process 1 events occur then the course of brain events could be strongly affected by mental effort. The timing of the Process-**1** events is, within the orthodox Copenhagen/von Neumann quantum theoretical framework, governed be the choice made by the experimenter/agent, and this choice is not specified by any known law of physics. But a rapid sequence of pairs of questions and answers (Process-**1**/Process-**3)** events can, by virtue of the quantum laws themselves, hold a particular pattern of neurological activity in place, against the physical forces that would, both in the absence of such pairs, and also in classical physics, tend quickly to disrupt it. If this pattern of neurological activity were to be a "Template for Action" (i.e., an "executive" pattern of neurological activity that tends to produce a specific action) then the prolongation of the activation of this "executive pattern" of brain activity can tend to cause the intended bodily action to occur, in accordance with William James's "ideo-motor" theory of action (James, 1890, p. 522). (According to that theory, it is the holding in place of the idea of an action that tends to make that action happen.)

This fact that a sufficiently rapid sequence of consciously selected probing events can hold the associated pattern of physical activity in place longer than what would be specified either by the classical laws of motion or its quantum analog, Process **2**, is *an automatic consequence of the quantum laws of motion*. It has been extensively studied by quantum physicists, both empirically and theoretically, under the title “The Quantum Zeno Effect.”

This quantum process can provide a physics-based account of the causal efficacy of conscious willful effort. This account corresponds closely to the ideas of William James, as is made evident by the following quotations:

> ``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.'' (James, 1890, p. 564)

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.''

The conclusion here is that the apparent capacity of our conscious efforts/choices to influence our physical actions, which seems so puzzling and necessarily illusory within classical physics has a straightforward explanation within quantum theory. This causal connection follows directly from the orthodox quantum laws of motion. Moreover, the details of how the process works is in amazingly close accord with William James’s account of how willful effort brings about intended actions. Unlike the situation in classical physics, these willful choices themselves are not controlled by the known laws of physics. There is, therefore, no warrant in contemporary physical theory for the assumption that our human choices are strict consequences of local mechanical processes akin to, or analogous, to those appearing in the classical physics approximation. The classical approximation completely wipes out the uncertainties within which the free choices are allowed to act. This approximation contracts the spreading cloud-like structures of quantum theory into narrow pencil-like beams, thus eliminating the freedom provided by quantum theory. In contrast to the Process **3** choices on the part of nature, which are subject to statistical laws, and hence are forced to be “random”, the Process **1** choices on the part of agents are not subject to any known law, statistical or otherwise, and hence need not be ruled by pure chance. Thus is important, because it is often thought by the ill-informed that all of the indeterminateness introduced by quantum theory is controlled by statistical laws, and is hence “random”. But the crucial Process 1 choices on the part of the agents are not subject to any known statistical or deterministic conditions.

In quantum theory the connection between mental effort and physical action can be explained as a *causal consequence of the laws of quantum physics*, combined with an assumption that an agent’s conscious effort to produce some experientially characterized effect increases the rapidity of a set of Process-**1** probing actions that focus attention on the intended experience. The experiential side of each such action/event is specified by an intended (projected) experiential state. The physical side collapses the prior physical state of the brain to a sum of two parts. The first part is the part of the prior state in which the neural correlate (Template for Action) of the conscious intention is definitely present. The second part is the part of the prior state in which the neural correlate of the conscious intention is definitely not present. In quantum theory there are generally parts of the prior state that are not compatible with either of those possibilities. Those parts are eliminated by Process **1**, which is thus associated with *asking* a question. Process **3** gives nature’s immediate answer: it collapses the state to the ‘Yes’ part or to the ‘No’ part. These pairs of abrupt events can be regarded as the “posing by agents” and the “answering by nature” of specific experientially formulated questions with ‘Yes’ or ‘No’ answers. These events with their experiential and physical sides are the basic realities of quantum mechanics. Between such event-pairs the state evolves via the local mechanical process **2**.

This tripartite quantum dynamics involving Choice, Causation, and Chance (Processes **1**, **2**, & **3**, respectively) and the implementation of Will (Volition) via the conscious control of the rapidity of Process **1** events, provides the mathematical and logical foundation of a pragmatic quantum approach to neuro-psychology. But how well does this quantum approach work in actual practice?

Pashler’s Analysis.

A great deal of experimental work in the field of The Psychology of Attention is summarized in Harold Pashler’s recent book of that title [Pashler, 1998].

Pashler organizes his discussion by separating perceptual processing from post-perceptual processing. The former covers processing that, first of all, identifies such basic properties of stimuli as location, color, loudness, and pitch, and, secondly, identifies stimuli in terms of categories of meaning. The post-perceptual process covers the tasks of producing motor and cognitive actions beyond mere categorical identification. Pashler emphasizes [p. 33] that ``the empirical findings of attention studies specifically argue for a distinction between perceptual limitations and more central limitations involved in thought and the planning of action.'' The existence of these two different processes, with different characteristics, is a principal theme of Pashler's book. [pp. 33, 263, 293, 317, 404.] He argues that the former processes are carried out in parallel, but that the latter processes, which seem to require effortful choosing, operate in series, and have a capacity that, although limited, can often be enlarged by willful effort.

Pashler's conclusion is based on the analysis of a huge array of recent experiments. But the central finding is succinctly illustrated in a finding dating from the nineteenth century, namely that mental exertion reduces the amount of physical force that a person can apply. He notes that ``This puzzling phenomena remains unexplained.'' [p. 387]. However, if we take the sequence of Process I events associated with an agent to have a limited "capacity" in terms of events per second, then this effect is a natural consequence of quantum theory. Creating a physical force by muscle contraction requires a *conscious effort* that prolongs the existence of the neural template for action, in opposition to the Process-**2**-generated tendency of the brain to evolve toward a more relaxed state. This prolongation is produced by the Quantum Zeno Effect, and its effect 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.

This example is just one simple case, but it illustrates the general principle. The identification of Pashler's limited central serial "capacity" with the rate of occurrence of Process **1** events, assumed to be increasable by willful effort, up to a limit, appears to explain the general features of all of the many diverse empirical results cited by Pashler in support of his thesis. (Schwartz, Stapp, & Beauregard, 2003; Stapp, 2001)

The apparent success of this quantum psychophysical theory in accounting for Pashler's data does not mean that classical physics could not be supplemented in some ad hoc way that would enable it to match that performance. However, the von Neumann theory allows the data to be explained directly in terms of *the already existing explicitly described tripartite process that constitutes the core of contemporary basic physical theory*, whereas an explanation based on classical physics is predicated on the untenable idea that the classical concepts of causation can be extrapolated from the motions of planets and falling apples to the motions of ions inside nerve terminals. It also rests on a theory that is not only demonstrably false, but claims to be dynamically and logically complete without entailing the existence of a part of reality that we know does exist, namely human consciousness. In contrast, von Neumann's equations specify definite dynamical connections between consciousness and brain activity, and they do so in a theoretical framework that automatically entails all of the valid predictions of classical physics. So what is the rationale, in neuro-psychology, for rejecting the fundamental equations of contemporary physics, which can straightforwardly account for the "observed" causal efficacy of consciousness, and also explain all of the valid classical features of phenomena, in favor of an extrapolation into a microscopic regime where it is known to fail of classical concepts that leave consciousness out?

The Libet Experiment.

Perhaps the best way to understand the essence of the quantum approach to consciousness is to see how it applies to the famous Libet experiments pertaining to willful action. (Libet, 2003)

The empirical fact established by the Libet data is that when an action is 'willed'– such as 'willing' a finger to rise– a readiness potential (RP) appears *before* the conscious experience of 'willing' appears. The most straightforward conclusion is that the causal efficacy of "free will" is an illusion. The motion of the finger seems clearly to be caused by neural activity that began well before the conscious act of "willing" occurs. Thus consciousness is seemingly a *consequence* of neural activity, not a *cause* of it.

The quantum mechanical analysis of this experiment leads to the opposite conclusion.

In the Libet experiment the original commitment by the subject to, say, "raise my finger within the next minute" will condition his brain to tend to produce *a sequence* of potential RP's distributed over the next minute. That is, the cloud of quantum possibilities will begin to generate a sequence of possible RP's, each one beginning at a different time. Each such RP will be associated with the 'Yes' answer to the question "Shall I choose (make an effort) to raise my finger now?" If the answer is 'No' then the 'template for the action of making an effort to raise the finger at that moment' will not be actualized, and the brain state associated with the answer 'No' will then evolve until the possibility of actualizing the template and RP corresponding to a later moment of choice arrives. When the brain activity associated with any *one* of these RP's reaches a certain triggering condition the Process **1** action associated with that particular RP will occur. Because the original commitment is spread over a minute the probability, for any individual RP in this sequence, for Nature's answer to be `Yes' will be small. Hence most of the possible RP's up to the one corresponding to some particular moment will *not be actualized*: they will be eliminated by the `No' answer on the part of Nature. But for one of these Process **1** events the associated Process **3** will deliver the answer "Yes," and the associated experience e will occur. Up to this point the conscious will has entered only via the original commitment to raise the finger sometime within the next minute. But in order to be efficacious the later experience e must contain an element of effort, which will cause the Process **1** associated with this experience (or a very similar one) to occur quickly again, and then again and again, thereby activating the Quantum Zeno Effect. This will cause the finger-raising template for action to be held in place, and the effect of this will be the issuing of the neural messages to the muscles that will cause the finger to rise.. Without this willful effort, which occurs in conjunction with the answer ‘Yes’, the sustained activation of the template for action will not occur and the finger will not rise. The willful effort causes the rapid repetition of the Process 1 action to occur. This holds the template in place, which causes the finger to rise. Thus the rising of the finger is caused, in the quantum formulation, by the willful effort, in concordance with the idea expressed by James (1892, 227)

> “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 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 introduces no new idea, it will deepen and prolong the stay in consciousness of innumerable ideas which else would fade more quickly away.”

Applications in Neuropsychology

This theory has been applied to Neuropsychology. (Oschner, Bunge, Gross, & Gabriel, 2002; Schwartz, Stapp, & Beauregard, 2003). In these studies human subjects are first instructed how to alter their mental reactions to emotionally-charged visual stimuli by adopting certain mental strategies. For example, the subjects are trained how to reduce their emotional reaction to a violent or sexual visual scene by cognitively re-evaluating the content; for example, by interpreting or contextualizing it in a different way. Their reactions to such stimuli are then studied using fMRI under differing choices of mental set. The brain scans reveal profoundly different patterns of response to the stimuli according to whether the subject does or does not apply the cognitive re-evaluation. Without cognitive re-evaluation the brain reaction is focused in the limbic system, whereas when cognitive re-evaluation is employed the focus shifts to pre-frontal regions. This demonstrates the powerful effect of cognitive choices upon brain functioning.

This effect is not surprising. But now that this apparent causal connection between conscious choices and ensuing brain behavior has been empirically demonstrated scientists must determine what theoretical framework is best suited to the description and analysis of this data, and of the host of similar data that will follow.

A key *empirical* input variable in these experiments is the conscious willful choice by the human subject about how he or she will (mentally) act. Von Neumann quantum theory provides a theoretical framework for analyzing these data in terms of such psychologically described input parameters. In this physics-based framework these experiential choices are taken to be primitive parts of the *cause* of the subsequent brain activity, rather than mechanically determined and purely epiphenomenal side-effects.

The classical approach contradicts the quantum principles on two counts.

On the one hand, the idea that the conscious choices are consequences of a local mechanical process described by the purely physical laws is exactly the idea that quantum theory *had to deny* in order bring consistently into the dynamical equations the needed *choices* of the "basis states" or "basis vectors" that are essential to the successful deduction of predictions about outcomes of experiments. These choices are not explainable in terms of the local mechanical process. The Heisenberg Uncertainty Principle blocks any attempt to derive definite Process 1 choices from the local mechanical laws of motion.

On the other hand, the classical notion that our conscious choices must be causally ineffectual not only lacks any basis in contemporary fundamental physical theory, but is contradicted by those principles. The *possibility* of influences via the quantum Zeno effect of Process 1 choices upon physically described properties means that conscious choices *can in principle* have physical effects.

Thus the idea that casual connection between the empirically described conscious choices and the empirically described subsequent brain activity is a *one-way* connection *from* the physical *to* the experiential is the *reverse* of what contemporary physical principles can explain. These principles allow us to explain causal connections from mind to brain, but not yet the connection from brain to mind.

In view of these considerations, it is scientifically unwarranted to insist on forcing our understanding of the mind-brain connection to conform to the seventeenth century ideas of Galileo and Newton. Forcing science to wear classical blinders in this area of research has led, over the past three hundred years, to tomes of analysis and polemics, but no resolution. However, the basic lesson learned in atomic physics may apply equally well to neuroscience: our conscious decisions are more satisfactory constituents of basic scientific theory than the concept of atomic-sized planets..

How, then, does von Neumann quantum theory apply to neuroscience?

The basic elements of von Neumann's theory are the experiences of conscious agents, and the neural correlates of those experiences, the NCC's. The fundamental building blocks of quantum theory are “action-events” each of which either (1) poses a question associated with some possible experience e, or (2) gives a ‘Yes’ answer to such a question, in which case it actualizes both this experience e (i.e., puts e into a stream of consciousness) and also reduces the prior physical state to the part of that state that is compatible with that experience e, or (3) gives a ‘No’ answer, in which case it eliminates the part of the prior state that is incompatible with experience e. The physical state actualized in conjunction with the ‘Yes’ answer is the neural correlate of the conscious experience e.

Each posed question contains a *projected* version of an intended or expected experiential feedback. (According to this stipulation even an unexpected feedback must be “posed” before it can be consciously received.) A typical feedback is an experiential confirmation that an intended action has occurred.

Empirical data are represented experientially. The receipt of such datum is represented by an event, or action, that “reduces” the prior physical state of the brain to a new physical state that contains the neural correlate of that experience ”e”, and contains no components that are incompatible with that experience. This focusing first on empirical data and its incorporation into the physically described state, and then later on the effect of this change in the physical state upon the predictions pertaining to the structure of future experiences is a *pragmatic* approach.

But how is the necessary connection between the experiential and physical regimes established? The answer is by trial and error empirical testing of the correspondence between the feeling of the conscious effort and the feeling of the experiential feedback. Every healthy alert infant is incessantly engaged in mapping out the correspondences between efforts and feedbacks, and he/she builds up over the course of time a repertoire of correspondences between the feel of the effort and the feel of the feedback. This is possible because different effortful choices have, according to the quantum equations, different physical consequences, which produce different experiential consequences. This whole process of learning would seem to depend crucially upon the causal efficacy of chosen willful efforts: if efforts have no actual consequences then how can learning occur, and the fruits of learning be obtained by appropriate effort.

The focus here has been on the theoretical foundations of pragmatic neuroscience. However, von Neumann’s shifting of the boundary between the observer and observee tends to shifts the theory in an ontological direction..

The essential difference between quantum theory and classical physics, both ontologically construed, is that the classical state of the universe represents a purported *material* realty, whereas the von Neumann quantum state of the universe represents a purported *informational* reality. This latter reality has certain matter-like features. It can be represented in terms of micro-local entities (local quantum fields) that *usually* evolve by direct interactions with their neighbors. But the von Neumann quantum state represents the collective knowledge of all agents, and it changes whenever the knowledge of any agent changes. Thus the “physical reality” represented by the quantum state has the idealike quality of a representation of an absolute or objective kind of knowledge. Like knowledge, its representation of faraway things can instantly change when we acquire here knowledge of something known to be correlated to the faraway things. Moreover, it represents possibilities, potentialities, and probabilities, all of which can be viewed as idea-like qualities.

If one shifts over to an explicitly ontological interpretation, the question immediately arises “What systems besides human beings are *agents*?. The extreme difficulty in acquiring scientific data that bears on this question is a strong reason for staying, at present, close to a pragmatic stance based on “our knowledge”: the sum of all human knowledge.

Its worth noting that everything said about the von Neumann theory is completely compatible with there being very strong interactions between the brain and its environment. The state S(t) of the brain is what is known as the statistical operator (reduced density matrix) corresponding to the brain. It is formed by averaging (tracing) over all non-brain degrees of freedom, and it automatically incorporates all of the decoherence effects arising from interactions with the environment.

Von Neumann’s theory provides a general physics-based psycho-physical framework for neuroscience. We now turn to some efforts to tie this structure to the detailed structure of the brain

----

Bohm, D. (1952). A suggested interpretation of quantum theory in terms of hidden variables. *Physical Review,* 85, 166-179.

Bohm, D. J. (1986). A new theory of the relationship of mind to matter. *The Journal of the American Society for Psychical Research,*
80, 113-135.

Bohm, D. J. (1990). A new theory of the relationship of mind to matter. *Philosophical Psychology*, 3, 271-286.

Bohm, D, & Hiley, D.J. (1993). *The Undivided Universe.* London and New York: 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.

Hagen, S., Hameroff, S, & Tuszynski, J. (2002). Physical Review E65, 061901-1 – 061901-11.

Hameroff, S. & Penrose, R. (1996). Orchestrated reduction of quantum coherence in brain microtubules: a model for consciousness. J. Consciousness Studies 3, 36-53.

Heisenberg, W. (1958a). The representation of Nature in contemporary physics. *Daedalus* 87, 95-108.

Heisenberg, W. (1958b). *Physics and Philosophy*, New York: Harper,

James, W. (1890) *The principles of psychology Vol. II*. New York: Dover.

James, W. (1892) *Psychology: The briefer course*. In *William James: Writings 1879-1899*. New York: Library of America (1992).

Libet, B (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioural & Brain Sciences*, 8, 529-566.

Libet, B. (2003). Cerebral physiology of conscious experience: Experimental Studies. In N. Osaka (Ed.), *Neural Basis of Consciousness*. [Advances in consciousness research series, 49] Amsterdam & New York: John Benjamins.

Ochsner, K.N. & Silvia A. Bunge, James J. Gross, and John D.Gabrieli (2002). Rethinking feelings: An fMRI study of the cognitive regulation of emotion. *J. Of Cognitive Neuroscience,* 14:8, 1215-1229.

Pashler, H. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.

Penrose, R. (1986). *The emperor's new mind*. New York: Oxford.

Penrose, R. (1994). *Shadows of the mind*. New York: Oxford.

Putnam, H. (1994). Review of Roger Penrose, Shadows of the Mind,

New York Times Book Review, November 20, p.7. Reprinted in AMS bulletin: www.ams.org/journals/bull/pre-1996data/199507/199507015.tex.html

Schwartz, J. & Begley, S. (2002). *The mind and the brain: neuroplasticity and the power of mental force*. New York: Harper-Collins.

Schwartz, J., Stapp, H. & Beauregard, M (2003). The volitional influence of the mind on the brain, with special reference to emotional self regulation. In M. Beauregard (Ed.), *Consciousness, Emotional Self-Regulation and the Brain*. [Advances in Consciousness Research Series]. Amsterdam & New York: John Benjamins.

Stapp, H. (1999). Attention, intention, and will in quantum physics. J. Consciousness Studies, 6, 143-164.

Stapp, H. (2001). Quantum theory and the role of mind in Nature. *Found. Phys*. 31, 1465-1499.

Tegmark, M. (2000). Importance of quantum decoherence in brain process. *Physical Review* E61, 4194-4206.

Von Neumann, J. (1955/1932). *Mathematical Foundations of Quantum Mechanics.* Princeton: Princeton University Press. (Translated by Robert T. Beyer from the 1932 German original, Mathematiche Grundlagen der Quantummechanik. Berlin: J. Springer)