May 15, 2000 LBNL-45229

A Quantum Theory of Mind. *

Henry P. Stapp
Lawrence Berkeley National Laboratory
University of California
Berkeley, California 94720

Abstract

A theory of consciousness is erected on von Neumann's formulation of quantum theory. More specifically, a theory of attention and mental effort is developed. This presentation is aimed at psychologists: it assumes no familiarity with quantum theory. The theory is highly structured, compared to theories based on classical concepts, because it rests heavily on restrictive quantum conditions on the form of the interaction between mind and brain. These restrictions invest the theory with significant explanatory power. Supporting empirical evidence comes both from the general observations of William James and from recent experimental findings pertaining to attention and mental effort.

*This work is supported in part by the Director, Office of Science, Office of High Energy and Nuclear Physics, Division of High Energy Physics, of the U.S. Department of Energy under Contract DE-AC03-76SF00098

1. Introduction.

Classical physical theory provides an unsatisfactory foundation for a theory of consciousness. The precepts of classical physics require all brain activity to be in principle completely determined by bottom-up local atomic processes alone, with the experiential events that form a person's stream of consciousness making no difference at all either to his bodily behaviour or to what is happening in his brain. This has two unsatisfactory consequences. It means that there is no reason for the incredibly elaborate, complex, and sensitive system of human consciousness ever to have evolved, and hence no way to understand the causes of its development. And it means that conscious experiences must be added to the theory ad hoc, simply because we know that they exist, rather than because they are logically necessary parts of a unified dynamical theory of the natural world. Consequently, the explanatory power of any theory of consciousness based on classical physical theory is inherently limited by this fact that in classical theory consciousness is a free-floating supernumerary having no fixed logical connections to the physical world.

The situation is completely reversed if quantum theory is used. Within the quantum framework a rational approach leads naturally to theory in which mind enters into brain process as a integral and logically necessary component of the basic dynamics. Thus the fundamental physical principles imposes stringent conditions both on the way consciousness arises from brain action, and on the way it acts back on the brain. The theory of attention and mental effort that emerges from this tightknit approach has impressive explanatory power. It explains both the general features of attention and efficacious mental effort described by William James at the end of the nineteenth century, as well as certain empirical findings that have remained mysterious and unexplained since that time, and also many hitherto unexplained features uncovered by recent empirical work.

It is well known that the founders of quantum theory introduced the observer and his observations into physical theory. This innovation constituted a sharp break from the ideas of classical physical theory, which placed the conscious experiences of human observers outside the part of nature dealt with by physical theory. But according to quantum philosophy, as articulated by Niels Bohr: “In our description of nature the purpose is not to disclose the real essence of phenomena but only to track down as far as possible relations between the multifold aspects of our experience.”[1]. In the words of Heisenberg: “The conception of objective reality of the elementary particles has evaporated not into the cloud of some new reality concept but into the transparent clarity of a mathematics that represents no longer the behaviour of the elementary particles but rather our knowledge of this behaviour.”[2]. Human experience and knowledge thus became identified as precisely the thing that basic physical theory is about. Moreover, the human experimenter plays a key dynamical role: a human intelligence standing outside the physical system, as it is represented by the mathematical machinery of quantum theory, chooses which aspect of nature will be probed. In Bohr’s words:

“The freedom of experimentation, presupposed in classical physics, is of course retained and corresponds to the free choice of experimental arrangements for which the mathematical structure of the quantum mechanical formalism offers the appropriate latitude.”[3]

This Copenhagen approach is ideally adapted to the practical applications of quantum theory to laboratory experiments. It enables quantum physicists to avoid getting entangled in metaphysical issues that are, for all practical purposes, irrelevant to these important applications. This Copenhagen quantum theory rests on the idea of dividing nature into two parts: “an observed system”, which is described by the mathematical machinery of quantum theory, and “an observing system”, which includes the human observers and their measuring devices. The latter are described in everyday language, refined by the concepts of classical physical theory, without, however, suggesting that classical physical theory really holds exactly for this part of the world: ordinary language is simply the language that we humans do in fact use to communicate to others what we have done, in setting up the experiments, and what we have learned from our observations.

The practical virtues of the Copenhagen approach stem directly from this division of the dynamically unified world into these two differently described parts. Though useful in practice, this tearing asunder of the unified physical world is, from the standpoint of principle, a disaster. However, this problem at the level of principle was solved by John von Neumann, who showed that the predictions of Copenhagen quantum theory could be obtained in principle by treating the entire physical world, including the bodies and brains of the human observers, in the mathematical language of quantum theory, and allowing the experiences of each observer to interact dynamically, according to quantum precepts, with the brain of that observer. This gives a highly constrained theory that is, at the same time, both a logically coherent way of formulating quantum theory and also a suitable physical basis for a theory of consciousness.

There is, however, one effect, well known to physicists, that seems at first to block any important contribution of quantum effects to the efficacy of consciousness. This effect, called environmental decoherence, is the strong tendency of a warm wet system, such as a human brain, to be effectively decomposed by interaction with its environment into a statistical mixture of quasi-classical states. This decomposition would appear, superficially, to eradicate all macroscopic quantum effects that might contribute to the efficacy of mental effort in the control of brain activity [4].

Consideration of this decoherence effect played, in fact, a crucial role in the development of the theory of consciousness described here. It forced the theory to rely, in order to achieve strong effects of mental effort upon behaviour, on the Quantum Zeno Effect. This is a well-known and well-studied quantum effect that is not weakened by environmental decoherence, and it is the basis of the theory of efficacious mental effort described here. The essential point, which will be explained in detail, is that focus of attention is maintained by mental effort via the physical mechanism of the Quantum Zeno Effect, and this connection explains many empirical features of the mind-brain connection.

This article is aimed at mainly psychologists. No knowledge of physics is assumed. Indeed, it may be beneficial not to know any physics, because quantum theory alters or reinterprets everything that classical physical theory says. Hence understanding classical physical physics creates prejudices that have to be overcome. I intend here to be both precise enough to satisfy quantum physicists yet clear and simple enough to reach readers with no knowledge of either classical or quantum physics. Once I have introduced from scratch a few simple ideas, I shall actually derive the Quantum Zeno Effect, and show how it works to make mental effort causally efficacious in the control of physical processes in a human brain. Then I shall compare the resulting properties to both to the observations of William James, and empirical results of the last half century, with emphasis on results about attention and effort obtained in the last two decades.

2. The Form of the Interaction Between Mind and Brain

This work is based on objectively interpreted von Neumann/Wigner quantum theory. I have argued elsewhere [5] that the evolving state S(t) of von Neumann/Wigner quantum theory can be construed to be our theoretical representation of an objectively existing and evolving informational structure that can properly be called "physical reality". von Neumann/Wigner quantum theory is essentially a theory of the interaction between this objective physical reality and subjective human experiences. The earlier Copenhagen version was essentially subjective, being, fundamentally, merely a set of rules that allow scientists to make predictions pertaining to connections between possible human experiences. It renounced the effort to understand the reality lying behind these rules. The immediate objective of the von Neumann theory was to show how these empirically validated Copenhagen predictions about connections between human experiences can be deduced from a rationally coherent conception of an objectively existing physical reality that interacts in specified ways with the subjective elements described in the Copenhagen formulation. Thus von Neumann and Wigner supplied, first, a putative description of the underlying physical reality that the Copenhagen approach tries to ignore, and, second, the form of the interaction between this objective physical reality and the experiential realities that are the basis of the Copenhagen approach.

I shall now describe the essentials of quantum theory, and of the connection between Copenhagen quantum theory and von Neumann/Wigner quantum theory.

Suppose there were just one particle in the universe. And suppose that, according to classical ideas, it must lie at one of N possible locations. Then one could represent the universe theoretically by imagining a long row of N boxes, each capable of holding a number, and then placing the number "one" in the box corresponding to the location where the particle lies, and a "zero" in every other box. If one wanted to represent statistical knowledge or information, then one could put into each box a number representing the "probability", or "statistical weight" associated with that possible location of the particle.

To pass to quantum theory, for this simple case, one would represent the state S(t) of the universe at time t, not by a 'row' of N boxes, each containing some number, but by an N-by-N 'square array' of boxes, each containing some number.

The "diagonal" elements of this matrix play a special role. The "diagonal" elements of an N-by-N matrix are the N numbers that lie in the boxes situated on the diagonal, i.e., in a box such that, for some number n, the box lies in row n and column n of the square array. The diagonal element that lies in row n and column n represents, in our simple example, the probability that the particle is in location n. The off-diagonal elements do not have this simple statistical interpretation, but they enter into the dynamics, and hence cannot be ignored. A matrix in which all of the off-diagonal entries are zero is called a 'diagonal matrix'.

For any matrix M, the 'number' in row i and column j is denoted by M_{ij}. These individual numbers are called the matrix elements of the matrix M.

The 'diagonal' elements of M are the numbers M_{ii}.

Matrices are generalizations of numbers, and share many of their properties. One can multiply a matrix a number: the rule is simply to multiply each matrix element of the matrix by that number. One can add two N-by-N matrices: the rule is simply to add together the two matrix elements that occupy the same position in the two arrays. Two N-by-N matrices, A and B, can be multiplied together to form the N-by-N matrix called AB. The rule for constructing the matrix AB is this: the matrix element of AB that lies in row i and column j is defined to be a sum of N numbers each of which is a product of two numbers. The specific rule is this:

AB_{ij} = The sum over all N possible values k of $A_{ik}B_{kj}$.

Notice that the values i and j that define the row and column of the matrix element AB_{ij} appear also on the right-hand side of this equation. The right-hand side contains also a variable k that runs, in successive terms, over all the integers from 1 to N.

This multiplication rule may seem arbitrary, but from a certain point of view it is the natural rule. Multiplication of matrices is important in both mathematics and classical mechanics, as well as in quantum mechanics.

This multiplication rule does not ensure that AB=BA. This dependence of the product of matrices on the order of the factors is important. But apart from this dependence on the *order* in which they appear, matrices can be treated like ordinary numbers. If one merely preserves the "order" in which the matrix factors are written down, one can do all of the usual operations of addition, subtraction, and multiplication. This will make it easy actually to derive the Quantum Zeno Effect from elementary "arithmetic", without writing down any numbers.

The term 'operator' is often used in place of 'matrix'. An 'operator' represents, in fact, a generalization that covers, for example, the case in which. the discrete integer indices i and j are replaced by continuous variables. I often use the word 'operator', rather than 'matrix', but the difference is not important here.

Quantum dynamics consists of a deterministic evolution of the state of the universe, punctuated by a sequence of sudden quantum events, or 'jumps'. The deterministic evolution between events is controlled by the Schroedinger equation, which is the deterministic quantum counterpart of the Newton/Maxwell equations of motion.

A key feature of Copenhagen quantum theory is that for each event an experimenter, who stands outside the system that is represented by the mathematical machinery, must choose a specific question: he decides which aspect of the quantum system will be probed by the observing system. This question is either itself a Yes-No question, or it can be decomposed into a finite set of Yes-No questions.

Each such Yes-No question is represented within the mathematical machinery by an associated 'projection operator' P. A projection operator is an operator (or matrix) that satisfies PP=P.

In our simple idealized example, where the universe consists of a single particle that is located at one of N possible locations, a possible question might be: Is the particle located at one of the first n locations?. The projection operator P corresponding to this question is a diagonal matrix that has a "one" in each of the first n boxes on the diagonal and a zero in every other box. One can easily verify, from the multiplication rule defined above, that PP=P, as required.

A basic rule of quantum mechanics asserts that if a question represented by a projection operator P is asked at time t, and the state of the physical system (or universe) just prior to time t is called $S(t-)$, and the state just after time t is called $S(t+)$, then

$$S(t+) = PS(t-)P$$

or

$$S(t+) = (1 - P)S(t-)(1 - P).$$

That is, the state will make a sudden "quantum jump" from $S(t-)$ to either $PS(t-)P$ or $(1 - P)S(t-)(1 - P)$.

What is the effect of the change from $S(t-)$ to $PS(t-)P$?

To find the answer one must perform the indicated matrix multiplications.

Suppose P is the matrix described above that corresponds to the question: "Is the particle located in one of the first n of the possible locations?" Then the product $PS(t-)P$ has a very simple form. It is the matrix obtained from $S(t-)$ by leaving undisturbed the numbers in those boxes that lie *both* in one of the first n rows *and* in one of the first n columns, but setting to zero the numbers in all other boxes. Thus $PS(t-)P$ is the 'projection' of $S(t-)$ into the n-by-n part of itself that corresponds to the answer 'Yes' to the question associated with P: all the numbers $S(t-)_{ij}$ with either i or j bigger than n are set to zero, but those both i and j less than or equal to n are left unchanged.

The symbol 1 in $(1 - P)$ stands for the matrix with all 'ones' on the diagonal. Thus $(1 - P)$ is the diagonal matrix with 'ones' in the *final* $N - n$ diagonal locations and zero elsewhere. Thus $(1-P)S(t-)(1-P)$ is the same as $S(t-)$ in the boxes that lie *both* in one of the final $(N - n)$ rows *and* one of the final $(N - n)$ columns, but has zeros in all other positions.

We thus see that the reduction of $S(t-)$ to $PS(t-)P$ reduces the state to the part of itself that is completely consistent with the answer 'Yes' to the question associated with P. The alternative reduction of $S(t-)$ to $(1 - P)S(t-)(1 - P)$ reduces the state to the part of itself that is completely consistent with the answer 'No' to the question associated with P.

To compare the theory to experiment one needs to extract real numbers from the theory. The basic operation that achieves this is the 'trace' operation. The trace of a matrix is just the sum of its diagonal elements:

Tr M = The sum over all values i of M_{ii}.

We are now in a position to state the basic interpretive law of quantum theory: If the state of the universe at the time immediately prior to time t is represented by $S(t-)$, and the question associated with P is asked at time t, then the state $S(t+)$ just after time t will be either $PS(t-)P$ or $(1 - P)S(t-)(1 - P)$, and:

1) The probability that $S(t+)$ will be $PS(t-)P$ is

$TrPS(t-)P/TrS(t-)$, and

2) The probability that $S(t+)$ will be $(1-P)S(t-)(1-P)$ is $Tr(1-P)S(t-)(1-P)/TrS(t-)$.

To understand these probability formulas recall that the diagonal elements of $S(t-)$ were the probabilities associated with the corresponding possible locations of the particle. That means that $TrS(t-)$ should be 'one': the sum of the probabilities should add up to 'one'. In this case the denominators in the above expression would be 'one'. However, it is advantageous to let the diagonal elements represent only the *relative* probabilities. Then the denominator is needed to get the overall normalization of the probabilities correct.

Recalling the form of $PS(t-)P$ we see that $TrPS(t-)P/TrS(t-)$ is just the sum, immediately prior to time t, of the probabilities associated with first n diagonal possible locations of the particle. It is, therefore, indeed, just the probability that the answer will be 'Yes' to the question, asked at time t, "Is the particle in one of the first n of its possible locations?" The alternative form $(1-P)S(t-)(1-P)$ corresponds to the answers 'No'.

I have been speaking, for simplicity, about the example of a single particle with N possible locations. But the general rules apply equally well if one replaces the single particle by entire physical universe, and allows the N possibilities to be N possible states of the universe.

To get an adequate relativistic theory of the mind-brain connection one needs to consider relativistic quantum field theory, and, in particular, quantum electro-dynamics, as discussed in reference 5. But the general logical structure is unaltered.

The general theory has four basic equations. The first defines the state of a subsystem. If S is the operator that represents the state of the universe and b is a subsystem of the universe then the state of b is defined to be

$$S_b = Tr_{-b}S, \tag{2.1}$$

where Tr_{-b} means the trace over all variables *except* those that characterize

b. In quantum electrodynamics the variables that characterize b are the local fields located in the body/brain of the person. The variables i and j can each be written as a *pair* of variables, where the first member of each pair refers to b and the second member refers to the rest of the universe. Then Tr_{-b} is constructed by setting the *second* members of the two pairs equal to each other, and then summing the equated variables over all N possible values. The first members of the two pairs are retained as variables, and hence $S_b = Tr_{-b}S$ is a matrix whose indices refer to the system b alone.

The second basic equation is von Neumann's process I. This process specifies the question that is asked, but not the answer. If a question is asked at time t, and $S(t-)$ represents the limit of $S(t')$ as t' approaches t from values less than t, then the state at time t is

$$S(t) = PS(t-)P + (1-P)S(t-)(1-P). \tag{2.2}$$

Here P is a projection operator (i.e., $P^2 = P$) that acts as the unit operator on all degrees of freedom *except* those associated with the body-brain b.

The third basic equation specifies the (Dirac) reduction. This reduction specifies the *answer* that nature gives:

$$S(t+) = PS(t) = PS(t-)P \text{ with probability } TrPS(t)/TrS(t), \tag{2.3a}$$

or,

$$S(t+) = (1-P)S(t) \text{ with probability } Tr(1-P)S(t)/TrS(t). \tag{2.3b}$$

Between jumps the state evolves according to:

$$S(t+\Delta t) = \exp(-iH\Delta t)S(t)\exp(+iH\Delta t). \tag{2.4}$$

Here H is the energy (or Hamiltonian) operator of quantum electrodynamics.

The reduction (2.3) corresponds to a certain increase in information It specifies one bit of information, Yes or No, and implants that information in the objective state $S(t)$ of the physical universe, by means of the reduction of $S(t-)$ to $PS(t-)P$ or to $(1-P)S(t-)(1-P)$. This objective state $S(t)$, with variable t, can thus be regarded as the evolving objective carrier of all of the bits of information generated by all of the reduction events that have occurred up until time t, minus the information that has been destroyed by the various reductions.

Information is normally conceived to be associated with an interpreting system. In Copenhagen quantum theory each reduction is associated with an increment in human knowledge, and the interpreting system is the brain and body of the observer. Thus in Copenhagen quantum theory the question in not "Is the particle in one of the first n of the possible locations?" It is rather: "Will I find the particle to be in one of the first n of the possible locations?" A probing question must be asked, and this question must be such that the two answers, 'Yes' or 'No', are experientially distinguishable.

In Copenhagen quantum theory the projection operator P associated with the question acts on the physical system that is being probed by the devices that the human observer is observing. But in von Neumann-Wigner quantum theory the projection operator P, which is—as in the Copenhagen approach—associated with a question about what the observer's experiences will be, acts on the degrees of freedom of the brain (and possibly the body) of that observer. The occurrence of the experience that corresponds to the answer 'Yes' is associated with the reduction of $S(t-)$ to $PS(t-)P$. This change reduces the state of the brain of the observer to a state that is compatible with this experience. This change is the interaction between mind and brain specified by von Neumann/Wigner quantum theory. It is this interaction that will be discussed and exploited in what follows.

One main point is that in order for the quantum dynamical process to proceed, a sequence of questions must be asked. Each question is associated with a quantum processor, which in the original von Neumann/Wigner formulation is a human being.

This anthropocentric bias of the theory is due to the fact that the immediate goal of the von Neumann/Wigner theory was to reproduce the predictions of Copenhagen quantum theory, and in the Copenhagen approach science is a strictly human endeavour based on human knowledge. The von Neumann approach allows us to relax this condition, and consider more general quantum processors and address the problem of the evolution of consciousness. But here I shall focus on the narrower issue of the connection between human minds and human brains.

3. The Macrolocal Process of Asking the Question

The projection operator P cannot be local: any point-like projection would inject infinite energy into the processor. Because the reduction of S(t) to P S(t)P is a nonlocal physical process it has no counterpart in classical dynamics: it is a new kind of element, relative to classical physical theory. But it is a key element in both Copenhagen and von Neumann/Wigner dynamics.

The projection operator is associated with a question asked by a person, or, more generally, by a quantum 'processor' that plays a dynamical role similar to a person.

Which question the person (or processor) will ask is not specified by the known rules of quantum theory. In Copenhagen quantum theory the choice of which question to ask—of which aspect of nature to probe— is left to the discretion of the experimenter, who is considered to stand outside the quantum-mechanically described universe. In von Neumann/Wigner quantum theory the body-brain of the human experimenter is brought into the quantum-mechanically described system, but the choice of the question he asks remains unspecified by the known quantum mechanical rules. Hence orthodox quantum theories do not specify how 'the question' is fixed or determined. This indeterminacy constitutes a causal gap that opens the door to the possibility that mind can influence brain in ways not determined by brain alone. The fact that neither Copenhagen nor von Neumann quantum theory specifies how these questions are determined means that orthodox quantum theories leave the choices of the question to be asked essentially free.

Because objective physical reality is informational in character, rather than material, it is not clear that our subjective knowings could not be, in some sense, composed out of the objective physical stuff, or be completely determined by the objective physical aspects of nature by rules not yet known to man. Maybe future developments will shed some light on such questions. But present-day physical theory, though it makes many sound predictions, does not answer these ontological and causality questions: the objectively-described and subjectively-described aspects aspects of nature appear in the theory as differently described elements that interact in accordance with certain dynamical rules.

The process of 'Asking the Question' has an important property. It is neither microlocal, nor globally nonlocal: it is *macro* local. It is a process that is, in a certain important sense, localized in the body-brain of the person who is asking the question.

To see this, suppose there are two participants (i.e., persons) located at time t in two well separated regions.

Suppose at time t the first one asks the question represented by P_1, and the second one asks the question represented by P_2. Then, because these two operators act at the same time on degrees of freedom that are localized in two well separated regions of space, the rules of relativistic quantum field theory demand that

$$P_1 P_2 = P_2 P_1 :$$

the two projection operators commute.

The following result is then easy to derive:

The probability $< P_1 >$ for outcome 'Yes' to the question associated with by P_1 is not affected by just asking the question associated with P_2.

PROOF: According to (2.2), just asking the question associated with P_2 reduces S to the form

$$[P_2 S P_2 + (1 - P_2) S (1 - P_2)].$$

But then, according to (2.3a), the probability that the question associated with P_1 has the answer 'Yes' is

$$< P_1 >= TrP_1[P_2SP_2 + (1 - P_2)S(1 - P_2)]/Tr[P_2SP_2 + (1 - P_2)S(1 - P_2)]$$

$$= TrP_1[P_2S + (1 - P_2)S]/Tr[P_2S + (1 - P_2)S]$$

$$= TrP_1S/TrS.$$

To derive the second line one proceeds as follows. The numerator in the first line breaks into two terms, the first of which is $TrP_1P_2SP_2$. The general identity $TrAB = TrBA$ is the used to move the factor $B = P_2$ to the left of P_1, giving $TrP_2P_1P_2S$. Then the equation $P_2P_1 = P_1P_2$, is used to move this factor P_2 to the right of P_1. giving $TrP_1P_2P_2S$. Then the property $P_2P_2 = P_2$ is used to obtain the first term in the numerator of the second line. The second term in the numerator of the first line gives, by means of a similar treatment, the second term in the numerator of the second line. The two terms in the denominator are treated in the same way. This give line two. Then a simple addition in the numerator, another one in the denominator, gives the third line.

The first line represents the probability of getting the answer 'Yes' to the question associated with P_1 under the condition that P_2 is asked, but no information about the answer (to that question in region 2) is included.

The third line represents what the probability of P_1 would have been if no question had been asked in region 2. Thus the computation shows that *just asking* the question in region 2 (i.e., asking the question, and adding the properly weighted contributions of the two possible answers) does not affect the probability of obtaining the answer 'Yes' to question P_1 in region 1.

In this sense, the process of 'just asking the question' associated with P_2 is *macro*local; it has no instantaneous effects outside region occupied by the body-brain b of participant 2.

However, *Answering* this question means keeping one of the two terms and leaving out the other. But in that case the cancellations leading to the macrolocality property derived above no longer occur. Thus *Answering* the question is not macrolocal! It is globally nonlocal.

4. The Quantum Zeno Effect

In this theory the main effect of mind on brain is via the Quantum Zeno Effect, which I shall now describe, and prove.

Suppose the initial state is PS(t)P, and that the question associated with P is repetitiously repeated. If these questions are posed at intervals Δt then equations (2.4) and (2.2) give

$$S(t+\Delta t) = P\exp(-iH\Delta t)PS(t)P\exp(+iH\Delta t)P$$
$$+(1-P)\exp(-iH\Delta t)PS(t)P\exp(+iH\Delta t)(1-P). \tag{4.1}$$

If Δt is small on the scale of the characteristic times for $P \leftrightarrow (1-P)$ transitions to occur then in the terms where $\exp(\pm iH\Delta t)$ is sandwiched between a P and a $(1-P)$ one can use the approximation to it, $1 \pm iH\Delta t$. But then the contribution from the unit term, 'one', drops out, by virtue of the equation $P(1-P) = (1-P)P = 0$. Thus only the single term $\pm iH\Delta t$ survives, Hence one sees that every term that contributes to the $P \leftrightarrow (1-P)$ transition occurs multiplied by at least *two* powers of Δt. But then these $P \leftrightarrow (1-P)$ transition terms drop out faster than the other terms when one decreases Δt by increasing the rapidity at which questions are asked.

This is the Quantum Zeno Effect: if rate of asking the questions is increased then the state tends to get trapped in the subspace, Yes or No, that it is already in,

This effect is sometimes facetiously called the "watched pot" effect. According to the old adage "A watched pot never boils", attending intently to whether the water is boiling prevents it from boiling. Similarly here, rapid-fire questioning effectively replaces the Hamiltonian H by $PHP + (1-P)H(1-P)$: the other two terms, which give the $P \leftrightarrow (1-P)$ transitions, tend to be suppressed by rapid-fire 'Askings'

This derivation shows that the Quantum Zeno Effect (QZE) is just as effective for a statistical mixture S(t) of quasi-classical states as for any other state: environmental decoherence does not eradicate or weaken this particular quantum effect, which arises basically from the reduction mechanism.

The only power given to the mind by this theory is the power to choose the questions P, and the times at which they are asked. And the only macroscopic dynamical effects of these choices thus far identified are the consequences produced by the Quantum Zeno Effect. This effect is merely to keep the brain activity focussed on a question for longer than it would stay focussed if quantum effects were not present. This focussing of attention is increased by increasing the rate at which the questions are asked. It is reasonable to suppose, or propose, that this rate can be increased by mental effort. Then mental effort can, by virtue of the Quantum Zeno Effect, keep attention focussed.

5. Explanatory Power

Does the theory described above explain anything?

Consider the following passage from "Psychology: The Briefer Course" by William James [7]. In the final section of the chapter on Attention he 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 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. The delay thus gained might not be more than a second in duration— but that second may be critical; for in the rising and falling considerations in the mind, where two associated systems of them are nearly in equilibrium it is often a matter of but a second more or less of attention at the outset, whether one system shall gain force to occupy the field and develop itself and exclude the other, or be excluded itself by the other. When developed it may make us act, and that act may seal our doom. When we come to the chapter on the Will we shall see that the whole drama of the voluntary life hinges on the attention, slightly more or slightly less, which rival motor ideas may receive. ..."

Posing a question is the act of attending. In the chapter on Will, in the section entitled "Volitional effort is effort of attention" [7] James 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."

von Neumann/Wigner quantum theory, with the quantum Zeno effect incorporated, seems to explains naturally the way in which, according to William James, human volition acts in our lives.

I shall tie down this connection by proposing a specific model.

In the present context the questions that a person can ask are supposed to be such that an affirmative answer corresponds to a possible conscious experience. Let it be assumed that a person with brain-body b has at each instant t a certain set of possible questions that he can ask. Let P_b mean $Tr_{-b}P$, and let $\{P_b\}(t)$ be the set of P_b's corresponding to the set of possible conscious experiences at time t. Let $P_b(t)$ be the P_b in $\{P_b\}(t)$ the maximizes $Tr_b P_b S_b(T)/Tr_b S_b(t)$. If the brain is a system whose job is to construct options for quantum events that can actualize control sequences that can produce actions suited to the circumstance in which the body-brain finds itself, then this $P_b(t)$ should be the P_b corresponding to the 'best' question: it is the P_b that is given maximal statistical weight by the evolving brain.

Let it further be supposed that at each time t the processor/person can either consent, or not consent, to asking the question associated with $P_b(t)$. To accommodate our intuitive feeling that mental 'effort' seems to focus attention I add the postulate that the rate of 'askings' can be increased by mental effort. By virtue of QZE, increasing the rate of 'askings' should tend to keep attention focussed.

This is a simple theory in which the effect of mind on brain is highly constrained: the only variables under mental control are "consent' and 'effort': everything else contributing to fixing the question asked has been ascribed to the brain. This brain computation is macrolocal. Because this computation involves P it can access and evaluate as a unit activities in different parts of the brain.

This simple quantum model fits James's descriptions very well.

6. Explanatory Power: More Recent Data

Much has happened in psychology since the time of William James. So one must ask: How well does the close correspondence between this quantum theory of consciousness and observations about attention and effort by William James hold up in the light of recent empirical work?

In this section I give an account of the data on attention and effort described and analyzed in Harold Pashler's 1998 Book, "The Psychology of Attention"[8]. Pashler treats James with great respect, but explains how more recent empirical work addresses issues that James was unable to deal with.

Pashler ties most of the empirical recent research on attention that he considers to its impact on various models of attention. He focusses on four different models:

1. Early selection theory (Broadbent's 1958 Filter theory.) All stimuli are processed to the point where certain physical attributes are represented,

and then a filtering occurs that determines which single stimulus is processed for "identification", one at a time. (p.14, 219)

2. Late selection theories. "Identification processing" proceeds unselectively without capacity limitation, then a later process selects which of the "identified" possibilities is selected for awareness.(p.16, 221)

3. Controlled Parallel Processing. The person chooses which stimuli are attended to, and the attended-to ones are processed, but the unattended-to ones are not fully processed. (p.224)

Pashler finds that evidence rules out possibilities 1. and 2. and favors 3. , which, however, "provides only the roughest scaffolding for a general account of perceptual attention."

Further evidence shows that when several stimuli are selected for attention they compete for a limited resource of identifying capacity.

Pashler claims that "attention" must, in scheme 3, be identified either with the gating part that selects what is to be attended to, or with the capacity part that controls the allocation of the identification-process resource. So one has two subpossibilities:

3.A. Attention controls which stimuli are selected for identification.

3.B. Attention controls allocation of limited identification-capacity.

Pashler asserts: "one thing we cannot choose is to maintain the presupposition inherent in commonsense talk and early and late selection theories that selectivity and capacity limitation are inherent features of the very same process or mechanism." (p.229)

Quantum theory makes one process both select stimuli and allocate a finite resource. This process is the ongoing process consisting of a sequence of reductions. Each reduction associated with an answer 'Yes' is associated with a conscious event. Hence the process in question is connected to conscious processing.

The reduction process effectively selects which stimuli will be actualized by picking out, via the reduction of S_b to $P_bS_bP_b$ (in case the answer is 'Yes'), from the morass of prior superposed partially processed stimuli, those stimuli that, by contributing to $Tr_b P_b S_b(t)$, caused the question to be asked. The reduction also controls, by actualizing the high-level control mechanism selected by $P_b S_b P_b$, the further brain processing produced by the high-level process actualized by the reduction event.

However, these reduction events are a limited resource that can be augmented by mental effort, to produce more of them per second, or divided by directing different events to different tasks.

This quantum theory appears to fit the "rough scaffolding" suggested by recent empirical studies, but seems closer to intuition that what Pashler deemed possible: the same mechanism of reduction both selects the stimuli that are actualized and fully processed, from a morass of stimuli that are processed in parallel superposed branches, and it also allocates the limited resource of the actualized higher-level processing.

I do not pretend that this sketch constitutes a full theory of attention and effort. But it arises naturally from the basic principles of physics, and already injects a fair amount of specificity into the mind-brain connection.

I shall now examine in more detail the various empirical findings described by Pashler that seem pertinent to this theory. They all seem to be in line with the theory.

I do not claim that these features could not be explained by a theory based on classical physical theory. But classical physical theory does not lead in any natural way to the idea that conscious experiences should even exist at all. Hence it would, on the one hand, appear to be intrinsically less constrained, hence less explanatory. On the other hand, it may be unable to explain, in any natural way, features that arise naturally in quantum theory from the macrolocal character of the reduction process, or from the fact that the quantum processing involves the whole state S_b, which is the quantum analog of the *whole statistical ensemble* of quasi-classical states, or from the natural quantum link between effort, rapidity of conscious events, and confinement of attention.

Pashler has a section "Mental Effort" in the chapter "Automaticity, Effort, and Control". He cites there first the 1968 experiments of Collyer that show that "incentives to perform especially well led subjects to improve both speed and accuracy." In the quantum model incentives should lead to mental effort, which would increase the rate of processing, and also increase accuracy insofar as the task is such that high-level processing improves accuracy.

Pashler says that it is interesting that the speedup was additive with variables affecting different processing stages, and that this additivity is not easy to interpret. But it would be expected if effort increases the overall rate of the conscious processing events, as it does in the quantum model. [I need to examine more closely the details of the experimental findings, but have not yet done so]

Pashler then considers a 1995 experiment by N.E. Johnson, Saccuzzo, and Larson that shows that incentives improve performance more for complex process that seem to require more conscious attention than for simple, more automatic tasks. This is in line with the QTOC (Quantum Theory of Consciousness), since effort merely speeds up the event frequency, which keeps attention focussed: this improves performance only on tasks where focussed attention is important.

Pashler then cites a 1983 study by Sanders which seems to be directly interpretable as evidence that effort tends to sustain focus of attention. A main basis of the QTOC is the physical explanation, on the basis of the fundamental principles of quantum physics, of how effort tends to sustain focus of attention.

Pashler then cites a 1986 work of Vidulich and Wickens, which shows that increasing the rate at which events occur in experimenter-paced tasks often increase effort ratings without affecting performance. This is to be expected from QTOC by virtue of the fact speeding up the task requires speeding up the mental processing, and effort is needed to do that: the non effect on performance should hold provided the task is simple enough so that good performance does not require an extended period of focussed attention.

A 1988 work of Yeh and Wickens shows that if the task difficulty is increased to a point that causes poor performance then effortfulness ratings generally increases. QTOC suggest that poor performance would lead to the need for a greater focus of conscious attention which would require more effort.

Vidulich and Wickens (1986) find that increasing incentives often increases workload ratings and performance. This would be expected from QTOC if the task is such that performance is enhanced by conscious focus of attention: incentives would then lead to effort to perform better by increasing focus of attention. [Again, a more detailed examination of the experimental findings is needed here]

The next section of Pashler's book is entitled "Subjective Effort: Causes and Consequences?"

He asks why people find effort taxing, and considers that people's aversion to exerting effort may be that it consumes energy, which from an evolutionary perspective, is better not squandered. However, the biophysical fact seems to be that mental effort does not markedly increase energy consumption.

Increasing the rapidity of mental events should not increase energy consumption significantly.

Pashler cites a puzzling phenomena discovered and confirmed at the end of the nineteenth century: "Almost all mental tasks reduce the maximum [physical] force [that one can exert], often by as much as 50 percent. This puzzling phenomena remains unexplained."

But according to QTOC the maximum effort increases to a maximum the number of brain events per second. This has the effect of maximally focussing attention on the attended task of producing force which has the effect of maximally controlling brain process in such a way as to tend to produce the maximum physical force. If attention is divided, then the action on the brain produced by the mental effort to produce the physical force is correspondingly reduced, and this reduced controlling action upon the brain will reduce the amount of physical force applied. This heretofore unexplained phenomena is thus an automatic consequence of QTOC.

This brief survey of the work on attention and effort described by Pashler, while still a work in progress, suggests that the natural quantum explanation of the features of conscious phenomena observed by Wm. James may extend to the more subtle features revealed by new data.

References

1. N. Bohr, "Atomic Theory and the Description of Nature", Cambridge, CUP. (1934)

2. W. Heisenberg, "The Representation of Nature in Contemporary Physics", Deadalus, **87**, 95-108 (1958)

3. N. Bohr, "Atomic Physics and Human Knowledge", Wiley, New York, (1958) p.72

4. Max Tegmark, "The Importance of Quantum Decoherence in Brain Process," Phys. Rev E, **61**, 4194-4206 (2000).

5. H.P. Stapp, "From Einstein Nonlocality to Von Neumann Reality," http://www-physics.lbl.gov/~stapp/stappfiles.html
quant-ph/0003064

6. H.P. Stapp, "Attention, Intention, and Will in Quantum Physics," in J. Consc. Studies **6**, 143-64 (1999).

7. Wm. James, "Psychology: The Briefer Course", ed. Gordon Allport, University of Notre Dame Press, Notre Dame, IN. Ch. 4 and Ch. 17

8. Harold Pashler, "The Psychology of Attention",
MIT Press, Cambridge MA (1998)