Consciousness and Quantum Theory: Strange Bedfellows

Barry Loewer
Rutgers University

> When I look at the scale of the apparatus I know what it reads. Those absurdly delicate, hopelessly inaccessible, global correlations obviously vanish when they connect up with me. Whether this is because consciousness is beyond the range of phenomena that quantum mechanics is capable of dealing with, or because it has infinitely many degrees of freedom or special super selection rules of its own, I would not presume to guess. But this is a puzzle about consciousness that should not get mixed up with efforts to understand quantum mechanics as a theory of subsystem correlations in the nonconscious world. (David Mermin 1998)

The nature of consciousness and the interpretation of quantum mechanics are two subjects that excite great interest. Even more exciting then is the idea percolating through certain quarters that there are deep and significant connections between the two. Among those who have advocated a quantum mechanics-consciousness connection are physicists Roger Penrose, Eugene Wigner and Henry Stapp, philosophers David Chalmers, Michael Lockwood, and Quentin Smith and even a judge, David Hodgson, and an anesthesiologist, Penrose's co-author Stuart Hameroff.[1] Why do these, and many of those who attend the huge consciousness conferences in Tucson, think that quantum theory has anything special to do with consciousness? There seem to be two kinds of reasons. One is that according to the standard way of thinking about quantum theory – aka "the Copenhagen Interpretation" – measurement and observation play a central role in physical reality in ways that are utterly different from classical mechanics. The theory's founding fathers said and current orthodoxy concurs that quantum mechanics requires for its *very formulation* reference to the measurement process; and while it might not be a majority view among physicists, it is often said that a measurement is not completed until it is registered in the mind of a conscious observer. Some physicists have taken this so far as to claim that reality is indeterminate until observed – or as it has been put "the moon is not there until someone looks." The other, complementary, reason is that the problem of understanding the relationship between consciousness and physical phenomena is *so hard.* Advocates of the quantum-consciousness connection think that it is so hard, in part, because the physical phenomena are understood in terms of classical physics. They suggest that progress can be made on this problem by recognizing that the physical basis of consciousness involves specifically quantum mechanical phenomena. Stapp, Penrose, and Smith, for example, claim that while classical mechanics is incapable of accounting for consciousness, quantum mechanics succeeds in providing explanations of how experience, unity of mind, free choice emerge from physical states.[2] Quantum mechanics-consciousness enthusiasts see a mutual need: quantum mechanics needs

[1]Lockwood, Chalmers, Penrose, Smith,

[2]Stapp

consciousness for its formulation – consciousness needs quantum mechanics as its physics. Thus mutual necessity makes for strange bedfellows.

Here I will be mainly concerned with the idea that quantum mechanics implicates consciousness although I will also make a few remarks about whether philosophers have much reason to look to quantum mechanics to illuminate philosophical issues concerning consciousness. But first, since this book is primarily a collection of papers on philosophical problems of consciousness, I will provide a quick tour of some of the main features of quantum theory that explain why it is thought to involve consciousness in some way.

1. **Quantum Mechanics**

Quantum mechanics is the framework for fundamental theories of microscopic systems and phenomena. There are quantum theories of various kinds of elementary particles, of molecular bonding, of the interactions between electromagnetic radiation and atoms, and of the forces that bind nucleons. And while there is not yet a satisfactory quantum theory of gravity, it is thought by most physicists that a true theory that includes gravity (along with the other forces) will also be a quantum mechanical theory. Since macroscopic physical systems (gases, starts, measuring devices, living organisms, etc.) are composed of quantum mechanical systems these too will conform to the laws of quantum theory even if their complexity prevents detailed application of quantum mechanics in explanations and predictions of their behaviors.

The main novel feature of quantum mechanics is its notion of *physical state*. In classical Newtonian mechanics the state of a (isolated) system of particles and fields at a time t is specified by specifying the positions and momenta at t of the particles and the values of the fields. The deterministic equations of motion describe how the state evolves. In quantum mechanics the state of a system is specified by a vector or *wavefunction* ψ in a Hilbert space (appropriate to that system). ψ contains all the information that can be known about S. Physical quantities – so called "observables" like position, momentum, energy, spin, field values etc. – are identified with certain *operators* O on the Hilbert space. The wavefunction corresponding to an observable possessing a specific value b1 is expressed in vector notation as $|O=b1\rangle$. The state represented by this wavefunction is said to be an "eigenstate" of O with value b1.

Quantum mechanics contains a vast amount of information concerning the quantum mechanical states of various kinds of microsystems, how to prepare systems in various states, and how to measure various observables systems. It also includes some basic propositions concerning the nature of states: how states connected with the values observables, how states evolve, and a rule for predicting the outcomes of measurements.[3] Among these are:

1) The superposition principle: for any collection of states (of a given system) there are possible states that corresponds to vector sums (or "superpositions") of the vectors corresponding to those states. For example,

[3] There are many places that philosophers can go to for a quick course on quantum mechanics and its philosophical problems. The best in my view is David Albert's *Quantum Mechanics and Experience* (Albert 1992).

the state c1|O=b1> + c2|O=b2> (where c1 and c2 are called the "amplitudes" of the states in the superposition.).

2) A physical quantity O has the value b for a system S if and *only if* the wavefunction ψ of S is an eigenvector of the operator associated with O. (This is called the "eigenvector-eigenvalue link")

3) The wavefunction ψ of a (isolated) system S evolves in conformity with a linear deterministic law (e.g., Schrödinger's equation).

4) If the state of S is ψ(S) = c1|O=b1> + c2|O=b2> +c3||O=b3> +.....and if an ideal measurement of O is made on S then the probability of obtaining result O=bi is ci squared (Born's Rule).

Ever since its inception quantum mechanics has been a theory in search of an *interpretation*. By "an interpretation" is meant an account of what a world would be like in which the theory is true. For a theory like quantum mechanics that aspires to be a theory of everything (or almost everything) physical that will involve an account reconciling quantum mechanics with our common sense conception of the macroscopic phenomena including measurement. An interpretation of a theory might involve additions or modifications either to the theory or to our common sense conceptions in order to achieve these goals. That quantum mechanics needs interpreting becomes obvious as soon as one notices is that there are observables O1 and O2 such that no quantum state is an eigenstate of both. In this case O1 and O2 are then said to be "complementary (and their associated operators "non-commuting"). Heisenberg's uncertainty principle – where the quantities are momentum and position – is an instance of complementarity. A state which is the superposition of eigenstates of O with different values is not an eigenstate of O but is an eigenstate of some other "complementary" observable.[4] But what does it mean for a particle, e.g. an electron, to possess a position but no determinate momentum or the other way round or, as is the general case, neither a determinate position nor a determinate momentum?

Matters get even more puzzling when one considers so-called "entangled states like the following state of a pair of electrons 1 and 2:

EPRB $\sqrt{1/2}$ (|up1>|down2> + |down1>|up2>)

'|up1>' refers to the state in which the x-spin of electron 1 has the value 'up' and so on. EPRB is not an eigenstate of the spins of the electrons in any direction but it is an eigenstate of the spins being correlated. Further, according to principle 3 the probability that a measurement of spin (in any direction) on particle 1 (or 2) will result in the value "up" (or "down") is ½. This is very puzzling. What can it mean for particles to have correlated spins but no determinate spins? And what does it mean that an x-spin measurement on either particle has a ½ probability of finding the electron with a determinate (either "up" or "down") spin? Einstein invoked a state similar to EPRB to argue that quantum theory is *incomplete*; that is there is

---

[4]Part of the rationale for the "only if" part of (1) which is responsible for this consequence is that it has proved *experimentally* impossible to devise a way of simultaneously measuring both O1 and O2. But even when ✈ is not an eigenstate of O (3) tells us the probability of outcomes of measuring O.

more to physical reality than the quantum state. In a nutshell his argument was that if two particles in an EPRB state are far apart and the spin of one is measured and a result e.g. ‘up’ is obtained then since the spins are correlated the spin of the other particle can be known to be ‘down.’ The measurement on one particle would appear to change the physical facts pertaining to the other. But this would involve a “spooky” non-local effect that seemed to Einstein to be incompatible with special relativity. To avoid this one apparently has to assume that both particles had determinate and correlated spins all along; in other words, that the quantum state isn’t a complete specification of the electron’s properties. Unfortunately, supplementing quantum mechanics with further, so called “hidden variables” is not an easy or straight forward matter. It can be shown that it is not possible for all observables to possess values that are disclosed by measurement and, as John Bell showed, the non-locality that bothered Einstein cannot be avoided.[5] Indeed it was and to an extent still is widely, although erroneously, believed that “hidden variable” theories are impossible.[6] I will briefly discuss one kind of hidden variable theory a bit later.

2. Quantum Theory Needs Consciousness

So what has all this to do with consciousness? To see why some physicists make a connection we need to discuss what quantum mechanics says happens when a measurement is made on a quantum system and when a human being observes that measurement outcome. Let O be an observable of a system S with two possible values b1 and b2 and M a measuring device with a macroscopic observable Q with three possible values Ready, B1 and B2. A measurement of O recorded by Q is an interaction in which the measuring device starts off in state |R> and if the state of S is |O=b1> then at the conclusion of the measurement the state of M+S is |P=B1>|O=b1> and if the state of S is |O=b2> then the state of M+S at the conclusion of the measurement is |P=B2>|O=b1>. It follows from (2) that if S starts out in a superposition |O=b1> + |O=b2> then at the conclusion of the measurement interaction the state of M+S is

MEAS |P=B1>|O=b1> + |P=B2>|O=b2>

MEAS is not an eigenstate of either O or P.[7] It follows from (1) that S does not possess a determinate value for O and also that M does not possess a determinate value for P! Schrodinger’s famous thought experiment concerns a cat which acts as a measuring device of an observable O of a quantum mechanical system S. The cat’s state of aliveness records the value of O. Schrodinger observed that it follows that if S’s state is a superposition of values of O then at the conclusion of the measurement the cat + S will be in a state like MEAS which is not an eigenstate of the Aliveness variable; i.e. the cat is not alive and is not dead but is in a

---

[5]The impossibility of supplementing the quantum mechanical state with hidden variables so that all observables possess values is shown by theorems is shown by Kochen-Specker and the inevitability of non-locality is demonstrated by Bell. For discussion of these points see Bub (1999), Bell (1987).

[6]It is mildly ironic that Bell’s theorem mentioned in footnote 5 is often cited as showing the impossibility of hidden variable theories while Bell was a proponent and developer of Bohm’s hidden variable theory.

[7]Although MEAS is an eigenstate of the values of O and B being correlated. That is why the interaction is a measurement.

superposition of the two conditions.[8] This consequence not only flies in the face of what we believe (that the cat will be determinately dead or determinately alive) but undercuts principle 3 that connects quantum mechanics with experiment. That principle says that the measurement in our example will result in one of the states |B1>|O=b1> or |P=B2>|O=b2> and not in the superposition MEAS.

To bring consciousness (finally) into the picture let's see what quantum mechanics says will happen if a human observer looks at a measuring device or cat in the state MEAS. I will suppose that there is a quantum state A1 (or collection of states) of an observer's brain corresponding to her experiencing seeing a live cat and an orthogonal state (or collection of states) corresponding to her experiencing seeing a dead cat. Then if the whole process is governed by Schrödinger's law the wavefunction of the system+cat+observer will evolve into the following state:

OBS) |M=M1>|P=B1>|O=b1> + |M=M2>|P=B2>|O=b2>

OBS is not an eigenstate of the observer consciously observing a live cat or of her consciously observing a dead cat. In OBS there is no determinate matter of fact concerning the observer's experience. It is important to keep in mind that a superposition of these two conscious states is not some other conscious state whose content is some vague amalgam of the two. It is a state with no determinate conscious content at all. One might think that states like OBS seldom arise, But if the evolution of state is governed by Schrödinger's law states like MEAS and OBS – that is, states that are not eigenstates of familiar quantities – will be the rule rather than the exception. In fact, as the wavefunction evolves it is very unlikely that it will ever be an eigenstate of there being *any conscious observers at all.* Obviously, this result undercuts quantum mechanics since it entails that we don't exist and never have any evidence for the theory.

We can see how considerations involving consciousness can enter into interpreting quantum mechanics. As physics has developed it has undermined certain aspects of our common sense view of the world; what Sellars (1963) calls 'the manifest image'.[9] For example, it appears that the moon and sun are approximately the same size, that earth stands still while the sun revolves about it; that ordinary objects (e.g. this table) are made of more or less homogeneous matter. But we know that these are just appearances and we have explanations based in physics (and psychology) of why things appear so but are actually quite different. If quantum mechanics understood realistically and including the universality of Schrodinger's law were true then almost no part of the manifest image would be correct; even that we have determinate appearances or even that we exist! Perhaps we can learn to live with the idea that elementary particles lack determinate position and even that cat's lack determinate *aliveness*; but the thought that our mental states always lack determinateness is

---

[8] If this is not problem enough what we have just described occurs in a measurement conflicts with what principle (4) says. According to it if the state of S is c1|O=b1> + c2|O=b2> then the probability that a measurement of O will obtain result b1 is c1. But, as we have just seen, if the measurement interaction is governed by (3) the state at the conclusion of the measurement is not one in which O has any determinate value at all!

[9] See Maudlin (1994) for a very nice discussion of the problem if interpreting quantum mechanics in terms of Sellars's 'manifest' and 'scientific' images.

epistemologically self-defeating. We cannot explain that our mental states *merely appear* to be determinate since that assumes that the appearance is determinate. We feel certain that if we were to look at a cat in MEAS our experience would have a determinate content. We would experience the cat as dead or experience the cat as alive *even if the cat is in a superposition of being alive and being dead.* A *minimal* requirement on any adequate interpretation of quantum mechanics is that the actual quantum state supports the existence of conscious observers with sufficiently many determinate conscious states to provide an evidential basis for quantum theory. I will call the requirement that any adequate fundamental theory of the world satisfies this the "determinate consciousness condition"(DCC). We also saw that there is a conflict between 1 and 2 on the one hand and 3 on the other. An adequate interpretation of quantum mechanics will have to reconcile this conflict as well as satisfying DCC. There are basically two strategies for interpreting quantum mechanics that have a reasonable chance of succeeding at this.[10] One is to modify the dynamical law (Schrödinger's equation) either by restricting its application or replacing it. The other is to reject the eigenstate-eigenvalue link.

The so called "orthodox solution" follows the first path. According to it the Schrödinger law is restricted to interactions that are *not* measurements.

2') Except for *measurements* the wavefunction ψ of a system S evolves in conformity with Schrödinger's law.

Measurement interactions satisfy the collapse postulate:

4) The collapse postulate: In an interaction in which an observable O of S is being *measured* then the state does not evolve linearly (as in 2) but rather "collapses" so that at the conclusion of the measurement it is an eigenstate of O.

The collapse postulate says that the post measurement state is not MEAS but one of the states $|P{=}B1\rangle|O{=}b1\rangle$ or $|P{=}B2\rangle|O{=}b2\rangle$ (with probabilities given by the squares of their amplitudes) which *are* eigenstates of O and P.

This 'solution' to the measurement problem raises the problem of saying exactly which interactions count as measurements. Not every interaction in which one observable

[10]There are also various *inadequate* responses to the measurement problem. Chief among these are Denial and Instrumentalism. There is a tradition of denying that Schrödinger evolution results in states that are not eigenstates of observables that record measurement outcomes because "there is no detectable difference" between a superposition of measurement effects like MEAS and a probabilistic "mixture" of outcomes when the measurement device interacts with the its environment. But this is simply wrong. As long as the wavefunction evolves linearly interaction with the environment will never result in a state that is an eigenstate of measurement outcomes. There is another tradition that says that quantum mechanics should be understood as a recipe for making predictions concerning the outcomes of experiments on micros-systems and shouldn't be used at all to describe macroscopic interactions like measurement. The trouble with this is that it is vague (what is "microscopic"?), puzzling (how is it that quantum mechanics describes microsystems but not macrosystems when the latter is composed of the former?), and unsatisfying (why does the quantum mechanics recipe work?).

becomes correlated with another counts as a measurement since there are some such interactions (involving microscopic systems) which are known to evolve in accord with (2) not (4). Some further condition must be placed on measurements. There are various proposals that have been made but the one that interests us gives a special role to consciousness.

A suggestion associated with von Neumann, Wigner, Wheeler and Stapp is that "measurements" occur only at the point at which a conscious observer interacts with the wavefunction.[11] I will call this "the consciousness collapse proposal" or CCP. Although the proposal is vague its advocates seem to have something like the following in mind. The quantum state evolves as described by Schrödinger's law until the systems it characterizes include a human being (or other system capable of consciousness) and can be represented as a superposition of eigenstates of states of consciousness. At that point the wavefunction collapses into one of the components of the superposition; i.e. into an eigenstate of consciousness. The probabilities of the various possible states are given by the collapse postulate.[12] So, for example the state of the electron and measuring device evolves into MEAS and then when observed by a human being the state of the electron and measuring device and human being evolves into OBS which instantaneously collapses into one of its components. The collapse, so to speak, propagates (instantaneously!) from the mind of the observer to the other systems with which it is entangled.

Henry Stapp (1995) characterizes the CCP this way:

> The key point, in the context of the mind/brain problem, is that this most orthodox interpretation of quantum theory brings the experiences of the human observers into the basic physical theory on at least a co-equal basis with the `physical' or `matter-like' aspects of the description: the matter-like aspects give only half of the dynamical and ontological story.

The metaphysical view underlying the CCP is probably best understood as a kind interactionist dualism. It is *dualist* in that it implies that the quantum mechanical description of the world is incomplete. A complete description would also specify which physical states are associated with consciousness. It is *interactionist* (as opposed to epiphenomenalist) in that consciousness has the causal power of initiating a collapse of the wavefunction. It is also a species of *emergentism* in that conscious states emerge from certain complex physical states (e.g. brain states).

One can see why the CCP would appeal to those philosophers who are already anti-physicalists or think that fundamental physics *in so far as it makes no reference to consciousness* is incomplete.[13] But as attractive as it may be to such philosophers it fails to

---

[11]More exactly the proposal is that when the quantum state a system containing a conscious observer evolves into a superposition of states some of which correspond to eigenstates of consciousness that wavefunction collapses to one of the states in the superposition with a probability given by Born's rule.

[12]Proposals of this sort are usually vague concerning exactly which mental states initiate collapse. Perceptions in which a person is conscious of a determinate outcome is one kind of mental event that initiates a collapse. Some (e.g. Stapp, Kane) suggest that acts of will (consciously making a decision) initiate a collapse (selecting one component of a superposition of states corresponding to different plans of action).

[13]For example Henry Stapp and David Chalmers take the CCP seriously.

provide a viable resolution of the measurement problem. There are two main problems; the one remediable and the other fatal. The remediable one is that it is not explicit about exactly when collapses occur. Are cats sufficiently conscious to collapse wavefunctions? Does any conscious state collapse a wavefunction or only those in which the observer is attending to a measurement outcome?[14] Until a clear criterion emerges for which states are the conscious states that initiate collapses we don't have an explicit theory. The fatal objection is this: given what we know about the early universe its quantum state was not an eigenstate of consciousness; i.e. there were not conscious observers during the first three minutes! But if the state of the early universe evolved by Schrödinger's law (as the CCP says it does) it would *never* result in a state that is an eigenstate associated with the existence of conscious observers. The problem is that the first collapse requires the existence of a conscious observer but the existence of a conscious observer requires prior collapses of states. The CCP can't get started.[15]

I can think of a modification of the CCP that might be thought to handle this problem. Suppose that whenever the state is a superposition of states that include states that support consciousness (i.e. states in which there are conscious subjects) that state collapses into one of its components. With this modification as soon as the state of the universe is such a superposition a collapse occurs and there is some chance that the collapse will produce a state in which there are conscious beings. The trouble with this suggestion is that it probably entails that collapses will occur too early in the history of the universe and too often. It is very likely that the state of the early universe was such a superposition.[16] If that is so then collapses would occur during the early universe at a rate contrary to what we know.[17]

The problem with the CCP is a problem for any version of the orthodox interpretation that identifies measurements with interactions involving systems of a certain degree of complexity. The problem is that it is unlikely that the initial state of the universe will evolve into an eigenstate of a complex measuring device. There is an ingenious collapse proposal that works quite differently due to Ghirardhi, Rimini, and Weber (GRW). Their proposal is, roughly, that the wavefunction of a system evolves in accordance with Schrodinger's law except that at any moment there is a chance that the wavefunction (as a function of particle position) is multiplied by a narrow Gaussian effectively collapsing it into a wavefunction almost all of whose amplitude is concentrated in a small region.[18] Such a wavefunction

[14] If any brain state that supports a conscious state initiates a collapse irrespective of whether the state's content is about a measurement then a person will be constantly collapsing states.
[15] Frank Arntzenius reminded me of this point. The objection is also fatal to any version of the collapse postulate which says that collapses occur only when the wavefunction is an eigenstate of the existence of a complex phenomenon (e.g. measurement).
[16] Recall that a state can be represented as a superposition in many (infinitely many) different basis vectors. Among these sets of basis vectors is a set of orthogonal vectors that are eigenstates of the physical quantities on which consciousness supervenes (i.e. these states specify what conscious beings exist, if any, and what conscious states they enjoy, if any). It is plausible that even the vacuum state (which, of course, is not an eigenstate of the quantities on which consciousness supervenes) is a superposition of such states some of which have positive amplitude for states that specify the existence of conscious beings.
[17] Collapses will be likely to localize particles and thus increase the velocity and temperatures of the system undergoing the collapse.
[18] See Ghriardhi et. al. (1986), Bell (1987), Albert and Loewer (1997).

corresponds to a localized particle. If a system consists of many particles – say 10^{23} – then although the chance in a small interval of time of any one of them undergoing a collapse is tiny the chance that at least one will can be practically 1. If the wavefunction of the system is one like MEAS in which the positions of many particles are correlated with each other then a collapse centered on one particle will have the effect of localizing the positions of all particles. GRW provides the same probabilities for experimental outcomes as orthodox quantum mechanics, at least as far as any experiments that are feasible.

GRW is much more plausible than the standard collapse proposal and its CCP version.[19] The concept of *measurement* doesn't occur in its fundamental law. Rather, it is a consequence of that law that those interactions we consider measurements have determinate outcomes.[20] Whether or not GRW satisfies the DCC depends on exactly what the physical basis – assuming there is one – of consciousness is; i.e. on what it supervenes. If mental states and processes supervene on the approximate positions of many entangled elementary particles (and the approximate values of fundamental fields correlated with particles) then – as for example in the firing of many neurons – GRW will succeed in making mental states determinate.[21]

The second strategy for dealing with the measurement problem is to give up the eigenstate-eigenvalue link. This is the approach taken by hidden variable theories and by so called "many worlds" interpretations. I will mostly discuss the latter since one of its versions involves conscious mental states in a fundamental way.

The many worlds interpretation was thought up by Hugh Everett in a Princeton dissertation he wrote under Wheeler the 1950s. For a few decades it was mostly ignored but now it seems to be replacing the Copenhagen account as the favorite interpretation among cosmologists and physicists who like to write about the interpretation of quantum theory. According to the many worlds interpretation a wavefunction like MEAS characterizes two kinds of 'worlds' in which each of the outcomes of the measurement is obtained. More generally, the whole universe is said to have a quantum state or wavefunction (the universal wavefunction) which as it evolves (in accordance with Schrodinger's law) describes a system of branching worlds. There is a problem of specifying exactly what worlds are described by the universal wavefunction. The problem is that a wavefunction can be represented in different basis sets. Exactly what worlds are described by a wavefunction depends on privileging a particular basis. If the basis is "approximate position" then it can be shown that given the kind of particles and fields that exist the branching worlds will be fairly well behaved. When an observable in a microscopic system interacts with the environment – as occurs in a measurement – the wavefunction branches into parts that develop independently of

[19]The objection we made against the CCP that the collapses would never take place doesn't hold against GRW since the collapses (multiplication by a narrow Gaussian) occur randomly even in the early universe.

[20]The reason is that measurement interactions involve systems consisting of many particles whose quantum states are correlated with each other and with the system being measured. Because the system consists of many particles there is a very high chance that at least one particle in the system will undergo a collapse and because the particles positions are correlated all the particles will end up in determinate positions.

[21]Not to say that GRW doesn't have its problems. For one, as I mentioned the collapsed states are never exactly eigenstates of position. For another, it is not settled exactly how to extend the idea to take into account all kinds of fields.

each other. It is claimed that from the point of view of an observer on a branch it appears that the wavefunction has collapsed and the world develops macroscopically in a more or less classical way.

Some find the many worlds interpretation too ontologically extravagant to be acceptable. Given the difficulty of coming up with any adequate interpretation I don't find this to be a serious worry. However, there are two more significant objections. One is that it might not satisfy the DCC. The other is that it provides no account of probability. Both of these worries have lead some theorists to put consciousness into the foundations of the theory. One way to guarantee that many worlds satisfies the DCC is to select as the basis that determines the branching worlds whatever physical variables consciousness supervenes on. This is basically the idea of the so called "many minds interpretation."[22] The worry about probability is that Schrödinger evolution is deterministic so there seems to be no way to ground or explain quantum mechanical probabilities.

Here is how the many minds theory is supposed to work . First, it assumes that there exists for any isolated (or approximately isolated) system (including the entire universe) a collection of quantum states – called "the mentality basis" for the system – each of which specifies exactly what minds would exist if that state were the quantum state of the system. If the state is a superposition of such quantum states then the minds associated with each element of the superposition all actually exist. So, for example, if we suppose that the states |M=M1>|P=B1>|O=b1> and |M=M2>|P=B2>|O=b2> belong to the mentality basis of observer+cat+quantum system and that the first is sufficient for the existence of a person seeing the cat alive and the second sufficient for the observer seeing a dead cat then there exists two minds associated with the observer; one seeing a live cat and the other seeing a dead cat.[23] On the many minds theory measurements have outcomes only in the minds. Thus the result of a measurement described by MEAS (where the measuring system does not involve a mind) is one in which neither outcome occurs (MEAS is a superposition of states corresponding to both outcomes) but once a sentient being becomes involved in the measurement and the resulting state is OBS mental states corresponding to seeing each outcome occur.

The worry concerning probabilities is that since the Schrödinger law is the sole dynamical law of the theory and is deterministic it is difficult to see how quantum mechanical probabilities can be accounted for by the interpretation. One idea is to suppose that associated with each branch in the consciousness basis are sets of minds and that quantum mechanical probabilities are values of a measure on these sets. Another suggestion is that the probabilities are dynamical chances governing the evolution of minds. It has been argued that the first fails to supply genuine probabilities and that the latter involves a very hard-to-believe dualism since it requires continuant minds whose identities over time fail to supervene on physical facts. In any case, the problem of probabilities is a difficult one for many worlds/minds

[22]The "many minds interpretation" is the brain child of Michael Lockwood (1989) and, simultaneously David Albert and myself (Albert and Loewer 1988). The former is an advocate of the view while the latter two authors attitude is better described as finding the view intriguing. There is also an important difference between Lockwood's account and the Albert and Loewer account concerning how to accommodate probabilities.

[23]The supposition is only to give a rough idea of how "many minds" works. The mentality basis is much more complicated and would determine a full set of mental states associated with a mind.

theories and whether it can be solved is a highly controversial matter.[24]

The ontology of the many-minds theory is the quantum state (which can be thought of as a kind of field in configuration space) *and* minds with mental states. It is assumed that the quantum state determines what and how many minds there are but the minds are not themselves part of the physical – i.e. quantum mechanical – reality. Thus the theory is committed to mental/physical dualism. However, there is a way of looking at the ontology which is inspired by a metaphysical view of Russell's so that it appears to be a monist theory with the quantum state and the mental state being two aspects of a single underlying reality.[25] According to Russell physics (and science more generally) reveals the causal and nomological structure of reality but is doesn't disclose its intrinsic nature. He went on to suggest that the intrinsic nature of reality is mental. Our knowledge of that intrinsic nature is restricted to the acquaintance that we have when we are conscious of our mental/brain states. However, the intrinsic natures of most properties are unknowable. The connection with quantum mechanics is the idea that the quantum state of a system characterizes its nomological/causal structure but that the intrinsic nature of the state corresponds to the conscious states associated with the quantum state. Of course, if by consciousness we mean state of a human or human like conscious observer then there are quantum states of the universe (and plausibly the state of the early universe was like this) which are eigenstates of there being no human conscious observers at all. Some proponents of the Russellian view (Lockwood and Chalmers) seem to suggest that such physical states – even states of isolated elementary particles – possess an intrinsic mental aspect. I have no idea what can be meant by this. Is there *something it is like* to be a muon? And if the answer to this is positive it is still completely mysterious how these so called "proto mental" states can compose to result in human consciousness; e.g. an experience of vertigo which most certainly there is something it is like to experience.

If the many minds interpretation were the only way of maintaining the Schrödinger law as the sole dynamical law then one might be willing to swallow its strange ontology of worlds and minds. But it is not. Another approach involves supplementing the state of a system with 'hidden variables'. This is the path taken by Bohmian mechanics.[26] According to (non-relativistic) Bohmian mechanics the physical state of a system consisting of n-particles is specified by a wavefunction and the positions of the particles. There are two dynamical laws; Schrödinger's law for the wavefunction and a deterministic law that specifies the particles velocities at t depending on the wavefunction at t. Probabilities enter into the theory by way of a probability distribution over the initial positions of the particles. Thus probabilities in Bohmian mechanics, as in statistical mechanics, reflect ignorance of the exact physical state. The most important features of Bohm's theory for the present discussion are a) that measurement, observation, and consciousness play *absolutely* no role in the formulation of the fundamental laws of the theory and b) since positions are always determinate all macroscopic entities and properties that supervene on position are determinate. For example, in MEAS the particle positions will be associated with one or the other of its components and so the measuring device (and Schrodinger's cat) will possess determinate positions (be determinately alive or determinately dead). Since it is a plausible assumption that our mental states supervene on positions of particles in our brain (and the wavefunction) it is plausible that Bohmian mechanics satisfies the DCC.

---

[24] For discussions see Lockwood (1989), Albert and Loewer (1988), Loewer (1996).

[25] Something along these lines is suggested by Lockwood (1989).

[26] For accounts of Bohm's theory see bohm and Hiley (1993); Albert (1992).

We have seen that quantum mechanics as embodied in principles 1-3 while on the one hand providing recipes that accurately predict the results of experiments involving elementary particles, atoms, molecules and so forth, on the other hand threatens to undermine our common sense beliefs in the existence of a determinate reality including our making determinate observations. It thus undermines itself. An interpretation of the theory must, at a minimum, fix things so that at least some of our experiences (enough to provide an evidential basis for the theory) are determinate. This is the DCC. It is via the DCC that considerations involving consciousness are properly brought into the interpretation of quantum mechanics. But there are proposals that go much further than satisfying DCC that make consciousness a more central player in the interpretation itself. The two main ones are the CCP and the many minds theory. We saw that the first suffers from vagueness and in the end fails at its task. The second is farfetched. If these kinds of approaches were the only ones that are capable of grounding an interpretation that solves the measurement problem and satisfies the DCC then perhaps it would be correct to say that quantum mechanics has put consciousness in the center of scientific ontology. But we saw that they are not. There are interpretation of quantum mechanics in which consciousness plays no special role and which plausibly succeed in resolving the measurement problem. The chief candidates are GRW and Bohmian mechanics. I do not mean to suggest that these accounts are free from puzzles and problems.[27] The interpretation of quantum mechanics as is fundamental physics itself is a far from finished matter.

3. Consciousness Needs Quantum Mechanics

I will conclude with a few reflections on the suggestion that consciousness needs quantum mechanics for its explanation. Of course, if quantum mechanics is the fundamental physical theory then if consciousness is a physical phenomena it is a quantum mechanical phenomena. But so are respiration, digestion, the weather, and every physical phenomena. It is very likely that neurophysiological states and events constitute and realize mental phenomena and it is certain that neurophysiological states and events involve chemical and electrical phenomena whose correct description involves quantum theory. A complete understanding of neurochemistry will involve quantum theory. But advocates of a quantum-consciousness connection mean to assert more than this. They think that the problems that philosophers have wrestled with concerning how physical phenomena can give rise to consciousness are intractable as long as one sticks with classical physics but that quantum theory can solve or help solve these problems. Here are some illustrative remarks:

> On the view that I am tentatively putting forward, consciousness would be some manifestation of this quantum entangled internal cytoskeletal state of its involvement in the interplay…between quantum and classical levels of activity”
>
> (Penrose 1994: 376)

The unity of a single mind can arise, in such a description, only if there is some form

[27]GRW has its ‘tails problem’ (see Albert and Loewer 1997) and problems reconciling it with special relativity and extending it to quantum field theory. Bohmian mechanics has problems concerning the interpretation of probability (see Loewer 2001) and extending it to field theory.

> of quantum coherence extending across at least an appreciable part of the entire brain" (ibid 372)

> The quantum framework leads naturally to the normal 'folk' concept of free will and personal responsibility. The key point is the concept of 'I'. In classical mechanics the personally experienced 'I' is not entailed by the (dynamically complete) physical principles, and it thus lies impotently, and hence without responsibility, outside the causal chain of physical events. In the quantum picture the experienced quality of 'I-ness' is experienced, and is therefore [part] of the stream of conscious events: the experienced 'I-ness' belongs to the experience, not vice versa. (Stapp 1995)

Penrose and Stapp think that when it comes to consciousness, functional, computational, and neurophysiological levels can be bypassed. We should go directly to physics – if that physics is quantum-mechanics – for an account of the nature of consciousness. I am very skeptical of such claims. Nothing of this sort has ever been proposed for other biological phenomena even though these too are constituted by quantum mechanical processes. Perhaps consciousness possesses certain special features and these involve quantum mechanics in a more substantial way. Well, I the first place we should be clear that quantum mechanics is of no help with what has come to be called 'the hard problem' of consciousness. This is the problem of explaining how the qualitative nature of experience – the *what it's like* aspect of consciousness arises out of the physical (neurophysiological) brain. There are arguments that purport to show that a *physical* duplicate of a person (which would of course also be a neurophysiological duplicate) may lack conscious experience. This conclusion is sometimes put as the claim that *zombies* are metaphysically possible. It has been shown that these arguments are unsound.[28] But the important point for my purposes is that they in no way depend on the specific nature of what counts as physical as long it is characterized in objective terms. However, the CCP does have consequences for physicalism and the 'hard problem'. If it were correct then a physical duplicate of the actual world that duplicated all its laws would also duplicate consciousness properties since consciousness is involved in the laws (the collapse law). So zombies (with respect to our world) would be impossible, but in a way quite different from what physicalists usually assume. Consciousness would be a fundamental causal feature of the world and not explicable in terms of objective physics (i.e. the wavefunction). The trouble with this is, as I have argued, that the CCP and akin proposals fail to solve the very problem that occasioned them.

There are some other features of consciousness that quantum mechanics has been claimed to illuminate. For example, it has been claimed (as in the quote from Stapp) that the apparent *unity* of mind can be accounted for in terms of the phenomena of quantum coherence. A state exhibits 'quantum coherence' if it is a relatively stable state of a relatively isolated system consisting of many particles that is a superposition of particle positions. The idea is that the entangled positions of the particles involve a kind of orchestration or unity of their evolutions. While it may be (though I think the evidence for this is slight) that there are states of neurons or microtubules that are coherent superpositions it is hard to see what this has to do with the unity of the mind. The unity of our mental lives involves our ability to integrate information from various sources, to store and retrieve memories, to form intentions, to conceive of ourselves as actors and so on. As far as I can see the claim that quantum

[28] Balog (1999).

coherence accounts for the *unity* of mind is wishful thinking founded on a pun.

It has also been suggested that quantum mechanics can save free will from the shackles of Newtonian determinism. An idea pursued by a number of authors is that in the course of making decisions the brain's quantum state evolves into a superposition of various alternative decisions and that making a choice involves that state collapsing into one of the alternatives.[29] Whether or not freedom of the will is compatible with underlying deterministic, or, for that matter, with indeterministic laws are big questions which I have discussed in some detail elsewhere.[30] Suffice it to say here that a quantum state collapsing into a particular choice rather than giving sense to the idea that the choice is made by the agent attributes it to an indeterministic process as much outside of the agent's control as it would be if the laws were deterministic. I have little hope that philosophers will be liberated from the problem of free will by quantum theory.

While it is understandable why it has become fashionable to connect a quantum mechanics and consciousness, news of their impending marriage is greatly exaggerated. Here I have mainly discussed the idea that quantum mechanics requires consciousness for its proper formulation. We saw that while there is a consideration – the DCC – that enters into evaluating the adequacy of interpretations of quantum mechanics, it is implausible that consciousness enters into physics in the ways suggested by quantum mechanics-consciousness enthusiasts. And while the nature of consciousness and its relation to neurophysiological and other physical phenomena are indeed hard problems there is little reason at present to think that quantum mechanics will have much to say about them.

---

[29] Nozick (1981); Kane (1996).

[30] Loewer (1998).