JEFFREY A. BARRETT

THE SINGLE-MIND AND MANY-MINDS V RSIONS OF QUANTUM MECHANICS

ABSTRACT. There is a long tradition of trying to find a satisfactory interpretation of Everett's relative-state formulation of quantum mechanics. Albert and Loewer recently described two new ways of reading Everett: one we will call the single-mind theory and the other the many-minds theory. I will briefly describe these theories and present some of their merits and problems. Since both are no-collapse theories, a significant merit is that they can take advantage of certain properties of the linear dynamics, which Everett apparently considered to be important, to constrain their statistical laws.

The standard theory of quantum mechanics has two dynamical laws.[1] The first says that a physical system evolves in a perfectly linear way whenever it is not being observed (for nonrelativistic quantum mechanics this is given by Schrödinger's time-dependent equation), and the second says that whenever an observation is made the object system nonlinearly jumps into a randomly determined state where the particular observation has a determinate outcome (this is the reduction or collapse postulate). The problem of reconciling these mathematically incompatible laws is known as the measurement problem. Everett proposed solving this problem by simply dropping the nonlinear collapse dynamics.

Everett's relative-state formulation of quantum mechanics is based on the assumption that "a wave function that obeys a linear wave equation everywhere and at all times supplies a complete mathematical model for every isolated physical system without exception" (Everett 1957, 316). He said that he intended to deduce the statistical predictions of the standard theory of quantum mechanics as "*subjective* appearances" (Everett 1973, p. 9). There is a long tradition of trying to figure out what Everett had in mind and what he ought to have had in mind. Albert and Loewer have recently described two new ways of reading Everett (See Albert and Loewer (1988) and (1989) and Albert (1992)). We will call these the single-mind and many-minds theories.

Erkenntnis **42**: 89–105, 1995.

2

Suppose we take Everett's proposal seriously and simply drop the nonlinear dynamics from quantum mechanics. Consider how one might describe a perfect x-spin measurement of a spin-1/2 system S initially in an eigenstate of z-spin.[2]

Suppose M is a perfect x-spin measuring device such that $|r\rangle_M|\uparrow\rangle_S$ (M ready to make an x-spin measurement and S in an x-spin up state) would evolve to $|\uparrow\rangle_M|\uparrow\rangle_S$ (M reporting x-spin up and S in an x-spin up state) and $|r\rangle_M|\downarrow\rangle_S$ (M ready to make an x-spin measurement and S in an x-spin down state) would evolve to $|\downarrow\rangle_M|\downarrow\rangle_S$ (M reporting x-spin down and S in an x-spin down state) – that is, suppose that the measurement interaction between M and S is such that M's pointer becomes perfectly correlated to the x-spin of S over the course of the interaction. It follows from this property and the linear dynamics that if the initial state $|\psi_0\rangle$ of $M + S$ is

$$|r\rangle_M \frac{1}{\sqrt{2}}(|\uparrow\rangle_S + |\downarrow\rangle_S) \tag{1}$$

the state $|\psi_1\rangle$ after M's x-spin measurement will be

$$\frac{1}{\sqrt{2}}(|\uparrow\rangle_M|\uparrow\rangle_S + \qquad {}_M|\downarrow\rangle_S). \tag{2}$$

In this nonseparable state neither M nor S have well-defined states of their own; rather, the system $M + S$ is in a superposition of M reporting $\uparrow$ and S being $\uparrow$ and M reporting $\downarrow$ and S being $\downarrow$. Consequently, if we asked M what the result of its x-spin measurement was, it would be in a superposition of reporting that it got x-spin up and x-spin down. This post-measurement state, however, has a curious property.

If we ask M what the result of its measurement was, it would be in a superposition of making mutually contradictory reports; but if we ask it whether it got a determinate, unambiguous x-spin result to its measurement, it would give us a determinate, unambiguous answer – 'Yes.'[3] Here's why. If the composite system $M + S$ were in the state $|\uparrow\rangle_M|\uparrow\rangle_S$, the answer to this question would be 'Yes,' and if $M + S$ were in the state $|\downarrow\rangle_M|\downarrow\rangle_S$, the answer would be 'Yes,' *so by the linearity of the dynamics*, if the composite system is in the state $|\psi_1\rangle$, the answer will be 'Yes.' Again, the linear dynamics along with the condition that M makes the correct report for the $|\uparrow\rangle_M|\uparrow\rangle_S$ and the $|\downarrow\rangle_M|\downarrow\rangle_S$ cases *requires* that it report that it got a determinate x-spin result if its state is $|\psi_1\rangle$. But consider what this report means. We have shown that M would report that it got a determinate x-spin result, either x-spin up or x-spin down, even though it is actually in a nonseparable superposition of recording mutually contradictory results. The standard interpretation of states tells us that a system has some property if and only if it is in an eigenstate of having that property. Since M is not in an eigenstate of recording one or the other x-spin result here, its report that it got a determinate x-spin result is simply false on the standard interpretation of states.

Quantum mechanics without the collapse postulate has several similarly curious properties that seem to have been particularly interesting to Everett.[4] The following is an informal catalogue of those properties that will prove useful later.

Determinate Result: After making a perfect measurement of any observable, M will be in an eigenstate of answering the question 'Did you get a determinate, unambiguous result?' with 'Yes.'

Repeatability: After a second perfect measurement of the same observable, M will be in an eigenstate of answering the question 'Did you get the same result for both measurements?' with 'Yes' if the object system is undisturbed between measurements.

3. Agreement: If another perfect measuring device N measures the same observable of the same object system and then communicates his result to M, both M and N will be in eigenstates of answering the question 'Did your result agree with the other measuring device's result?' with 'Yes' if the object system is undisturbed between measurements.

Relative Frequency: If M makes perfect measurements of the same observable on each of an infinite sequence of systems, all in the same initial state, M will approach an eigenstate of answering the question 'Were your results distributed with the usual quantum relative frequencies?' with 'Yes' as the number of observations gets large.

5. Randomness: If M makes perfect measurements of the same observable on each of an infinite sequence of systems, all

in the same initial state, M will approach an eigenstate of answering the question 'Were your results randomly distributed?' with 'Yes' as the number of observations gets large.

On the standard interpretation of states, all of these reports would in fact be false; nonetheless, these properties tell us what a good measuring device *would necessarily report* after a perfect measurement (or series of measurements) in any version of quantum mechanics that (1) takes the linear dynamics to correctly describe the time-evolution of every physical system and (2) predicts that if an observer is in an eigenstate of making some report, then the observer makes that report. Since Albert and Loewer's single-mind and many-minds theories satisfy both of these conditions, the above properties of the linear dynamics tell us what a good observer would report concerning his experiences in various specific situations – that is, if one of the above sets of antecedent conditions is ever met, the linear dynamics will require the observers to make the corresponding reports.

3

The single-mind theory might be thought of as an attempt to take the linear dynamics as a complete and accurate description of the time-evolution of the *physical* world yet ensure that observers always end up with determinate measurement results. Albert and Loewer do this by postulating the existence of *nonphysical* minds that always have determinate mental states and whose evolution is given by a mental dynamics.

One might reasonably want some motivation for such a move. Consider the problem again. If we take the linear dynamics to be universally true as Everett suggests, then an observer's physical state after an x-spin measurement will be something like

$$\alpha|\uparrow\rangle_M|\uparrow\rangle_S + \beta|\downarrow\rangle_M|\downarrow\rangle_S, \qquad (3)$$

which describes the observer as being in a nonseparable superposition of recording mutually contradictory results. An empirically adequate theory would presumably describe P as getting one or the other of the two possible x-spin results, but there is nothing in P's quantum-mechanical state that picks out a determinate x-spin result. Indeed, since the standard interpretation of states tells us that a system has a determinate property if and only if it is in a eigenstate of having the property, this physical state describes P as having no determinate measurement result since P is not in an eigenstate of getting x-spin up and it is not in an eigenstate of getting x-spin down. This suggests that we add something to our description of P that *does* describe him (or her) as getting a determinate measurement result and tells us what the result is – that is, if we want to keep the linear dynamics, then we presumably need to supplement the usual quantum-mechanical description of the observer with a parameter that determines his measurement result. The parameter is the truth-maker for propositions concerning the results of measurements. Albert and Loewer suggest that we supplement the usual quantum-mechanical state with the mental states of observers. This new parameter clearly determines an observer's measurement results, but it also suggests a commitment concerning the relationship between mental and physical states that one might not like. On the other hand, if we want keep the linear dynamics as an accurate description of the time evolution of all physical systems but also feel that we need to add something to the thcory that is not determined by the quantum state in order to describe an observer as getting a determinate result, then we are apparently committed to some sort of dualism (though not necessarily the mental-physical dualism suggested by Albert and Loewer).

Let P be an observer, and let $B_P(n, x)$ represent that P believes that the result of measurement n was x. The single-mind theory requires a *partial correspondence* between physical and mental states. It is assumed that there is generally a physical state $|(n, x)\rangle_P$ such that $|(n, x)\rangle_P$ implies $B_P(n, x)$. But it is also an important feature of the theory that $B_P(n, x)$ does not necessarily imply $|(n, x)\rangle_P$. It is this second feature that allows P's *physical state* to be a complicated superposition of physical states corresponding to mutually contradictory beliefs concerning the outcome of a specific measurement while P's *mental state* is always one where he has a determinate belief corresponding to one element of the superposition.

If the complete physical state $|(n, x) \ldots\rangle$ *implies* $B_P(n, x)$, then we will say that $|(n, x) \ldots\rangle$ is a measurement n belief eigenstate for P. Allowing i to vary over the possible outcomes of measurement n, the vectors $|(n, i) \ldots\rangle$ form an orthonormal set (since if P is in a physical state where he would report that the result was x_1, then the amplitude of the component of his state corresponding to the report that the result was x_2 must be zero if $x_1 \neq x_2$). This set might be completed to form an orthonormal basis for the Hilbert space used to represent the complete physical state. Choose any such basis, and call it a *measurement n belief basis*. When the physical state is written in this basis, each term describes P as either having a determinate belief or no belief at all concerning the result of measurement n. Along these lines one might define an eigenstate of full belief for P to be any physical state that would completely determine P's mental state.

The single-mind theory stipulates that while the deterministic linear dynamics always describes the time-evolution of the complete physical state, the time-evolution of an observer's mental state is probabilistic. Suppose that after some measurement n the physical state is $|\psi_n\rangle$. The probability that P's mental state ends up $B_P(n,x)$ is $|\langle(n,x)\ldots|\psi_n\rangle|^2$. In other words, to determine the probability of P's mental state being $B_P(n,x)$, write the post-measurement physical state in any of P's measurement n belief bases, then interpret the norm squared of the coefficient on the term $|(n,x)\ldots\rangle$ as the probability of P's mental state being $B_P(n,x)$. Note that this probability is determined solely by the post-measurement physical state, or as Albert and Loewer put it, "The probability that the mind will end up in a particular state is completely determined by the physical state of the observer + system measured" (1988, pp. 205–6). It turns out that Albert and Loewer actually want this probability to be determined by the post-measurement physical state *and the premeasurement mental state of the observer*, but let's keep the story simple for now and return to the details of the mental dynamics later.

An observer P who begins an experiment in an eigenstate of being ready to make an x-spin measurement of a system S in an eigenstate of z-spin, for example, would end up with his brain in a superposition of belief eigenstates corresponding to mutually incompatible results. The physical state of $P + S$ might be represented by

$$|\psi_1\rangle_{P+S} \quad \frac{1}{\sqrt{2}}[|(1,\uparrow)\rangle_P|\uparrow\rangle_S + |(1,\downarrow)\rangle_P|\downarrow\rangle_S] \tag{4}$$

Here one component of the physical state describes P as believing that the result was x-spin up and the other describes him as believing that the result was x-spin down. P's mind, however, would end up associated with only one of the two possible beliefs – in this case, there would be a 50% chance that *P*'s post-measurement mental state would be $B_P(n, \uparrow)$ and a 50% chance that it would be $B_P(n, \downarrow)$. The complete state is given by the quantum-mechanical state of the physical world together with the mental states of every observer.

Albert and Loewer are ultimately dissatisfied with the single-mind theory. Their primary worry is that it does not generally allow mental states to supervene on physical states – that is, a complete description of the physical world would generally fail to determine the mental state of an observer. The physical state $|\psi_1\rangle_{P+S}$, for example, is consistent with either $B_P(1, \uparrow)$ or $B_P(1, \downarrow)$. Albert and Loewer describe this type of non-physicalism as "especially pernicious," and they tell us that is it this lack of mental supervenience that leads them to consider the many-minds theory (Albert and Loewer 1988, 206).

The many-minds theory asks us to suppose that "every sentient physical system, every observer, has associated with it not a single mind but rather an infinite set of minds" (Albert and Loewer 1988, 206). Let $\mathcal{P}$ be the set of minds associated with observer *P*, let χ_n be the set of minds with a mental state where $B_P(n, x)$, and let μ be a measure on $\mathcal{P}$ such that $\mu(\chi_n)$ equals the norm squared of the coefficient on the term $|(n, x) \ldots\rangle$ when the physical state is written out in one of *P*'s belief bases. Albert and Loewer interpret $\mu(\chi_n)$ as the *measure* of *P*'s minds with a mental state where $B_P(n, x)$. In other words, one can determine the distribution of mental states of an observer's minds by expanding his physical state in one of his belief bases then associating a set of minds with a measure equal to the norm squared of the term's coefficient with each of the terms that describe him as having a determinate mental state. The state of each mind is then described by the term with which it is associated. The time-evolution of the mental state of each of an observer's minds is probabilistic, where the probability that the post-measurement state of a particular one of an observer's minds being correctly described by $B(n, x)$ is $|\langle(n, x) \ldots|\psi_n\rangle|^2$. On the other hand, the observer's "global mental state" is given by the measure μ. Albert and Loewer argue that since μ is determined by the quantum-mechanical state, the observer's global mental state supervenes on his physical state and that consequently the time-evolution of *this* mental state *is* deterministic.

The individual minds, as on the [single-mind theory], are not quantum mechanical systems; they are never in superpositions. This is what is meant by saying that they are non-

> physical. The time evolution of each of the minds on the [many-minds theory] is, just as on the [single mind theory], probabilistic. However, unlike the [single-mind theory], there are enough minds associated with the brain initially so that minds will end up associated with each of the elements of the final superposition. An infinity of minds is required since a measurement or a sequence of measurements may have an infinite number of outcomes. Furthermore, although the evolution of individual minds is probabilistic, the evolution of the set of minds associated with [a particular observer] is deterministic since the evolution of the measurement process is deterministic and we can read off from the final state the proportions of the minds in various mental states. (Albert and Loewer 1988, p. 207)

Consider an x-spin measurement again. The observer begins in an eigenstate of being ready to make an x-spin measurement of a system in an eigenstate of z-spin and ends up with a physical state that describes his brain as being in a superposition of belief eigenstates corresponding to mutually incompatible beliefs. One of these states describes him as believing that the result was x-spin up, and the other describes him as believing that the result was x-spin down. On the many-minds theory, all of his minds have determinate beliefs concerning the result of his observation, but not the same determinate beliefs. Here measure-one of the observer's minds would begin in mental states with the belief that he is ready to make a measurement, and with probability one, half of the observer's continuous infinity of minds would end up believing that the result was x-spin up and half would end up believing that the result was x-spin down.

Albert and Loewer argue that the many-minds theory has several advantages over other interpretations of Everett and other versions of quantum mechanics generally. The many-minds theory is true to Everett's fundamental idea that the time-evolution of the entire universe and every physical system is completely and accurately given by the linear dynamics: "There is no need to postulate collapses or splits or any other non-quantum mechanical *physical* phenomena".[5] The many-minds theory is "in accord with our very deep conviction that mental states never superpose". It "entails that the choice of basis vectors in terms of which the state of the world is expressed has no physical significance".[6] Also unlike the many-worlds interpretation, the many-minds theory encounters no special problems interpreting probability: "Probabilities are completely objective, although they do not refer to physical events but always to sequences of states of individual minds" (Albert and Loewer 1988, p. 208). Finally, its dynamical laws can be expressed in a local, Lorentz-invariant form, which means that the many-minds theory meshes well with relativity (Albert and Loewer 1988, pp. 209–10).

There is, however, another virtue that might be added to this list: the mental dynamics is strongly constrained by the properties of the linear dynamics mentioned in Section 2. The many-minds theory has two dynamical laws. The linear dynamics describes the time-evolution of the physical world, and the mental dynamics describes how the observer's minds evolve given the evolution of his physical state. At first glance, the mental dynamics looks ad hoc – it looks like an arbitrary rule cooked up just to make the theory consistent with our actual quantum-mechanical observations. It turns out, however, that the linear dynamics does not allow one much of a choice for the mental dynamics – that is, the evolution of the physical state in the many-minds theory strongly constrains the mental dynamics independently of specific empirical considerations.

The many-minds theory stipulates that an observer's physical state always evolves according to the linear dynamics. It follows then from the relative frequency and randomness properties described in Section 2 that if a many-minds observer measures the same observable on each of an infinite sequence of systems all in the same initial state, he would approach an eigenstate of reporting that his results were randomly distributed with the standard relative frequencies in the limit as the number of observations gets large. If one requires that *measure one* of the observer's minds end up with beliefs consistent with this report in the limit, then this strongly constrains how the observer's minds might evolve from measurement to measurement. If one further requires every length n sequence of measurement results to correspond to a possible mental state of each mind after n measurements and if one requires the mental dynamics to be trial-independent, then the post-measurement state of a particular mind must be randomly determined with the usual quantum probabilities, which is just what the many-minds theory says.[7] The basic idea here is that the mental dynamics looks much less ad hoc than it might because the linear dynamics tells us that an observer would report the usual statistics in the limit and we have reasons that are independent of specific empirical considerations for supposing that an observer's mental state is generally compatible with his reports.

4

Albert and Loewer consider the many-minds theory to have a decided advantage over the single-mind theory because the many-minds theory allows an observer's global mental state to be uniquely fixed by his physical state. As they describe the deal, "We have purchased supervenience of the mental on the physical at the cost of postulating an infinity of minds associated with each sentient being" (Albert and Loewer 1988, p. 207).

It is not quite right, however, to say that an observer's global mental state supervenes on his physical state. The mental state of each of the observer's minds is a random function of his physical state and independent of the states of his other minds. This means that the observer's global mental state *almost always* evolves deterministically. Likewise, his global mental state *almost always* supervenes on his physical state.[8] This lack of strict supervenience does not seem to be a very serious problem, but the *type of mental supervenience* that the many-minds theory provides is puzzling.

If one wants mental supervenience, one presumably wants the mental state that one is capable of introspecting right now, the mental state that one has epistemic access to, to supervene on one's physical state. I believe that I have a more-or-less definite mental state characterized by a single set of more-or-less consistent beliefs. But the many-minds theory tells me that I am associated with an infinite set of minds that most likely have wildly contradictory beliefs and whose mental states I cannot generally know. What comfort is it supposed to give me that my global mental state supervenes on my physical state when I don't even know what my global mental state is? This is made especially puzzling by the fact that neither my physical state nor my global mental state determine the state of the only mind I do know.

In order to get an observer's global mental state to supervene on his physical state to the extent that it does, his global mental state is characterized by the *measure* of his minds with each possible mental state, not by a description of which minds have which mental states. This means that there are an infinite number of different ways to assign what might be called *local* mental states to an observer's minds that would all correspond to the observer having the same *global* mental state. In other words, the global mental state associated with an observer fails to determine the local mental state of any of his minds.

Even if I had an infinity of minds, I identify so strongly with one mind that I wouldn't much care whether or not the infinity of minds together in some way supervened on my physical state. Being counter-intuitive is not a fatal flaw in a physical theory, but one might eventually decide that associating each observer with an infinity of minds costs more than it's worth.

While mental states do not supervene on physical states in the single-mind theory, an observer has only one mind. I think that one might consequently have good reasons for liking some version of the single-mind theory better than Albert and Loewer's many-minds theory.

Let's return to the single-mind theory and consider its mental dynamics in more detail. Suppose that an observer P makes an x-spin measurement of a system S initially in an eigenstate of z-spin and ends up with the physical state

$$|\psi_1\rangle_{P+S} = \frac{1}{\sqrt{2}}[|(1, \uparrow)\rangle\rangle_P|\uparrow\rangle_S + |(1, \downarrow)\rangle\rangle_P|\quad\rangle_S] \tag{5}$$

and a mental state where $B_P(1, \uparrow)$. Now what does the single-mind theory predict for the result of a second x-spin measurement when S is undisturbed between measurements? The linear dynamics tells us that if he repeats his measurement he will end up with the physical state

$$|\psi_2\rangle_{P+S} = \frac{1}{\sqrt{2}}[|(1, \uparrow)(2, \uparrow)\rangle\rangle_P|\uparrow\rangle_S + |(1, \downarrow)(2 \quad)\rangle\rangle_P|\quad\rangle_S], \tag{6}$$

but what would his mental state be? Let's suppose that probabilities are completely determined by the physical state alone and see what happens – call this the *primitive* single-mind theory. This is not what Albert and Loewer want. They want the mental state of an observer to depend on his past mental state in such a way that he would generally end up with the same result for a repeat spin measurement (Albert 1992, p. 127). The primitive single-mind theory, however, predicts that P would have a 50% chance of getting result $B_P(2, \uparrow)$ and a 50% chance of getting result $B_P(2, \downarrow)$. But if this were true, it would be possible for P to perform the second measurement and end up with a mental state where $B_P(2, \downarrow)$, which by hypothesis flatly contradicts the result of his first measurement, and we (presumably) know from experience that we always get the same result whenever we repeat a spin measurement on an undisturbed system.

At first glance this may look like a serious problem, but it turns out that there is a sense in which it doesn't much matter whether P's second result agrees with the result that P in fact got for his first measurement. This is because if P does get a different result for his second measurement, then he will not correctly remember the first result: if P ends up with a mental state where $B_P(2, \downarrow)$ after his second measurement, he will also have a mental state where $B_P(1, \downarrow)$ regardless of the fact that his mental state after the first measurement was formerly one where $B_P(1, \uparrow)$. This is because P will now be associated with the *first term* in the above superposition. So even if the mental dynamics were a random statistical function of the observer's current physical state and if his current mental state was thus largely independent of past mental states and if his actual experience was consequently pathologically discontinuous, he would not notice. Indeed, P could do no experiment that would determine what his former mental states actually were. The repeatability property of the linear dynamics tells us that he would report that a record of his first result agrees with his second result. This is true regardless of whether or not his first result was, as a matter of historical fact, what he now believes that it was. The properties of the linear dynamics and the relationship between mental states and physical states conspire to make it impossible for an observer to tell that his experiences were in fact pathologically discontinuous. The observer would remember his past experiences as a uniform and coherent sequence of events and he would take his physical records to support his memories, but his memories would generally be false and his belief that his physical records agree with his memories might best be described as an illusion. But this explanation itself suggests a problem.

If an observer's current mental state is determined independently of his past mental states, if his current mental state is simply a function of his current physical state, then even one's beliefs concerning one's own past experience would generally be unreliable, so how could one ever have reliable empirical evidence that would justify accepting the theory? At every instant an observer would have determinate beliefs concerning his past experience and he might even judge that the theory accounts for these beliefs; but if the theory were true, relatively few of the beliefs would be true. One would never be able to test the predictions of the theory against one's *actual* empirical experience; rather, one would at best only be able to determine whether one's *current* beliefs concerning one's empirical experience, which the theory itself tells us would generally be false, are compatible with the predictions of the theory. This apparently means that one could never have empirical evidence for accepting the mental dynamics specified by the theory. Unless a theory generally allows for reliable records of past experiences, one could have no empirical evidence for claiming that the theory correctly describes one's *actual experience*. Our usual notion of empirical adequacy presupposes a theory that allows for reliable records of past experiences – we will say that such a theory is *empirically coherent*. If a theory is not empirically coherent, then I cannot see how one could have empirical grounds for accepting the theory – in particular, one could have no grounds for judging whether the dynamics specified by the theory was in fact the right dynamics.

There is a way that the properties of the linear dynamics described in Section 2 might be used to constrain the mental dynamics so that we end up with an empirically coherent reformulation of the single-mind theory.[9] The argument has a transcendental flavor. We will start by simply assuming that the correct physical theory is empirically coherent. More specifically, we will suppose that the mental dynamics is such that an ideal observer's reports and beliefs concerning his own experience and other observers' experience (given their reports) are generally true. Because of the properties of the linear dynamics, this principle strongly constrains the mental dynamics.

Suppose our ideal observer makes a perfect x-spin measurement of an object system initially in an eigenstate of z-spin and ends up in the physical state predicted by the linear dynamics. The determinate result property tells us that the observer will report that he got either x-spin up or x-spin down as the result of his measurement. By stipulating that an observer's reports concerning his experiences are generally true, our new single-mind theory would require that the observer actually does end up with a mental state corresponding to *one or the other* of the two possible x-spin results. This might be taken as something of a justification for the single-mind theory's dualism – by associating him with a non-quantum-mechanical mind, we can make the observer's report concerning his own mental state true.

Suppose the observer got the result x-spin up to his first measurement, and then without disturbing the object system, he makes a second perfect x-spin measurement. The repeatability property tells us that the observer will report that he got the same result for both measurements. In order for his beliefs concerning his own experience to be true here, the new single-mind theory would have to predict that the observer gets x-spin up as his second result. Note that this places an important constraint on the mental dynamics – once a particular type of measurement is made on some system by an observer, the observer will generally get the same result for all future measurements as long as the system is not disturbed, which is just what Albert and Loewer want.

Suppose now that two observers are ready to make perfect x-spin measurements. The first observer makes an x-spin measurement on an object system that is in an eigenstate of z-spin and gets the result x-spin up; the second observer then makes an x-spin measurement on the same object system, which has not been disturbed since the first measurement, and gets a result. The agreement property tells us that if the two observers compare their x-spin results, they will conclude that their results agree, which means that if one wants an observer to be able to talk to his friends and routinely end up with true beliefs concerning their mental states, the new single-mind theory would have to tell us that the second observer also gets the result x-spin up. In other words, the mental dynamics is such that once an observer gets a measurement result, other observers would generally get the same result as long as the system is not disturbed.

As with the many-minds theory, the linear dynamics also places strong constraints on the statistical properties of the new single-mind theory's mental dynamics. Suppose that an observer performs an experiment where he measures the same observable of an infinite number of systems in identical states. The relative frequency and randomness properties require that the observer's physical state approach a state where he would report that his measurement results were randomly distributed with the usual quantum relative frequencies. In order for this report to be true with probability one, the mental dynamics would have to be such that the observer's measurement results really were randomly distributed with the appropriate relative frequencies, which obviously places very strong constraints on the mental dynamics. Specifically, if one assumes that the mental dynamics is trial-independent, in order for the observer to have a probability-one chance of his mental state agreeing with the report that we know he will make, it must be the case that each measurement result will be randomly determined by probabilities equal to the limiting relative frequencies.[10]

There is a more general conclusion. Any no-collapse theory that predicts that an observer will make a report if he is in an eigenstate of making the report must also predict that an observer's measurement results will be randomly distributed with the usual quantum-mechanical probabilities if we want a probability-one guarantee that the observer's statistical reports concerning the distribution of his results will be true of his actual experience in the limit.

6

Quantum mechanics without the collapse postulate has several suggestive properties that Everett apparently thought were important. These suggestive properties tell us what an ideal observer would report after a measurement in certain situations in any no-collapse formulation of quantum mechanics that predicts that an observer makes some report if he is in an eigenstate of making that report. But what exactly is the relevance of these properties to a satisfactory no-collapse formulation of quantum mechanics?

There is a role for the relative frequency and randomness properties in Albert and Loewer's many-minds reading of Everett. The linear dynamics tells us what an observer would report concerning his own experience in the limit as he preforms an infinite number of observations. If we want the report to be true of the beliefs of measure-one of the observer's minds in the limit and if we want a trial-invariant law, then we have strongly constrained the mental dynamics. The many-minds theory does have several advantages over other readings of Everett, but one might ultimately conclude that the advantages that the many-minds theory has over the single-mind theory fail to compensate for the disadvantages of associating every observer with a continuous infinity of minds.

One might consequently end up liking the single-mind theory better than the many-minds theory. But the primitive single-mind theory fails to be empirically coherent – it tells us that even one's beliefs concerning one's own past experience are generally unreliable, which means that if the primitive single-mind theory were true, one could have no empirical evidence for accepting it, at least not in the usual sense. On the other hand, if we presuppose empirical coherency, if we suppose that the mental dynamics is such that an ideal observer's reports and beliefs concerning his own experiences and other observers' experiences might generally be true, then the linear dynamics constrains the mental dynamics in a way that suggests a reformulation of the single-mind theory that might prove empirically adequate in a perfectly ordinary way. Further, this new formulation of the single-mind theory would exploit those properties of quantum mechanics without the collapse postulate that Everett apparently considered important.[11]

[1] What I am calling the standard theory is described by von Neumann (1932). This paper presupposes a basic understanding of how quantum mechanics works and the measurement problem. One might want to at least read Albert and Loewer (1988) before reading this paper.

[2] The following is a condensed version of a story told by Albert (1992, pp. 116–9). x-spin and z-spin are non-commuting quantum observables; that is, on the standard interpretation of states if a system is in an eigenstate of z-spin and thus has a determinate z-spin, then it has no determinate x-spin. Spin is a particularly simple observable: an x-spin measurement of a spin-1/2 particle can only result in spin-up or spin-down. All of the states in the paper are written in terms of x-spin eigenstates.

[3] Note that asking the measuring device M a question amounts to nothing more than measuring a particular physical observable of the composite system containing M.

[4] At least half of Everett's 1957 paper might be read as a discussion of such properties, but in the end it is unclear exactly how he took such properties to be relevant to his relative-state formulation of quantum mechanics. The first four of the following properties, especially the relative frequency property, are discussed in several places in the literature since Everett's 1957 paper. For a recent discussion see Albert (1992, pp. 116–23), and for a more detailed account see Barrett (1995).

[5] The reason that *physical* is emphasized here, of course, is that the many-minds theory supposes the existence of *non-physical* systems, the observer's minds, whose states are probabilistically correlated with the physical state of the observer's brain.

[6] This is not to say that the many-minds theory does not have a preferred basis – an observer's belief basis acts as a preferred basis of a sort. While one might say that this basis has no *physical* significance, it does have *empirical* significance. Specifically, the law that describes the relationship between an observer's mental state and his physical state explicitly appeals to the observer's belief basis to determine the proportion of his minds in a specified mental state given his physical state. Albert and Loewer's preferred basis, however, is arguably not a part of their *physical* theory – they take it to be something like an accidental property of the relationship between brain states and mental states, which is supposed to be a fact independent of one's physical theory.

[7] The many-minds theory predicts that *measure one* of the observer's minds will believe that his results were randomly distributed with the usual relative frequencies. Note that there is no trial-invariant statistical law that predicts that *all* of the observer's minds will have beliefs consistent with his report.

[8] That the many-minds theory does not *guarantee* mental supervenience was first noted by Marc Albert.

[9] The following version of the single-mind theory is incomplete since there are several properties of the mental dynamics that are unspecified. It is meant to show how a theory might take advantage of the suggestive properties of the linear dynamics by presupposing empirical coherency.

[10] It is important to note, however, that if an observer makes a report whenever he is in an eigenstate of making a report, then it is impossible to make *all* of his reports concerning his own experience true. An observer who measures the x-spin of systems initially in the same spin state would approach an eigenstate of reporting that he got a particular sequence of x-spin results, but he would also approach an eigenstate of reporting that each particular sequence of x-spin results fails to describe his experience. While there are logical systems where both reports might be taken to be true (systems that allow for supervaluation), standard first-order logic requires at least one of these reports to be false. The single-mind theory tells us that exactly one of these reports is false.

[11] I would like to thank David Albert and Barry Loewer for their comments on an earlier draft of this paper and for several discussions.

REFERENCES

Albert, D. Z.: 1992, *Quantum Mechanics and Experience*, Harvard University Press, Cambridge.

Albert, D. Z. and B. Loewer: 1988, 'Interpreting the Many Worlds Interpretation', *Synthese* 77, 195–213.

Albert, D. Z. and B. Loewer: 1989, 'Two No-Collapse Interpretations of Quantum Theory', *Nous* 23, 169–186.

Barrett, J. A.: 1994, 'The Suggestive Properties of Quantum Mechanics without the Collapse Postulate', *Erkenntnis* 41, 233–252.

DeWitt, B. S. and N. Graham (eds.): 1973, *The Many-Worlds Interpretation of Quantum Mechanics*, Princeton University Press, Princeton.

Everett, H., III.: 1957, '"Relative State" Formulation of Quantum Mechanics', *Reviews of Modern Physics* 29, 454–462; reprinted in Wheeler and Zurek (eds.) 1983.

Everett, H., III.: 1973, 'Theory of the Universal Wave Function', in DeWitt and Graham (eds.) 1973.

von Neumann, J.: 1932, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin; reprinted and translated as *Mathematical Foundations of Quantum Mechanics*, translated by R. Beyer, Princeton University Press, Princeton, 1955.

Wheeler, J. A. and W. H. Zurek (eds.): 1983, *Quantum Theory and Measurement*, Princeton University Press, Princeton.

Manuscript submitted September 1, 1993
Final version received March 16, 1994

Department of Philosophy
University of California
Irvine, CA 92717-4555
U.S.A.