The many-worlds interpretation of quantum mechanics: psychological versus physical bases for the multiplicity of "worlds".

Howard Barnum*

Los Alamos National Laboratory, MS B256, Los Alamos, NM 87545

Preface

This hitherto unpublished 1990 paper presents my views on the "many-worlds," or, as I prefer to call it, following Everett, the "relative state" interpretation of quantum mechanics. It was rejected by one journal in 1990, primarily on the (I think incorrect) grounds that it was too similar to the Albert and Loewer "many minds" version. It has seen limited circulation since then. My views are quite similar to ones since put forth by Simon Saunders, and subscribed to by David Wallace, among others. I believe that, within the relative state interpretation, when an observer interacts with a system in a measurement process in such a way that the combination of the two evolves into a superposition of the observer seeing macroscopically different apparatus pointer readings (or other macroscopic states), the observer's consciousness "forks" or "splits," because only those subspaces of Hilbert space describing definite macroscopic measurement results satisfy psychological/philosophical conditions for supporting a unified consciousness persisting through time. These conditions probably include things like the persistence of memory, and are probably closely linked to the "decoherence" of certain subspace decompositions or approximate decompositions, corresponding to persisting macroscopic "classical" structures that can support such persisting memories—relating the approach to the decoherence-based ones of, for example, Zurek, and Zeh, though this relation was not much stressed in the paper.

I still agree with most of what is in the paper, with some notable exceptions. First, the discussion of how one might have empirical evidence for "nonprobabilistic indeterminacy" seems shaky: the evidence cited might just as well be taken to be evidence "against independent trials," or "against exchangeability." (This does not vitiate the point that relative state-er's may need to establish not just that *if* there are probabilities they must be the Born ones, but that (on the relative state interpretation) the indeterminism facing a quantum experimenter about what will be subjectively experienced as an outcome is correctly viewed as probabilistic at all.) Second, as I think the 1990 referee pointed out, the "improvement on Everett's derivation" of probabilities (sec. 8.2), while it uses Gleason's theorem to avoid Everett's assumption that the probability of a wavefunction component can depend only on its squared modulus, does not justify the noncontextuality assumption implicit in Gleason's theorem. There are also more minor exceptions—for instance, I now see somewhat less value in de Broglie-Bohm-Bell nonlocal hidden variable theories than suggested toward the end of this paper.

Note that while I think it has much to teach us about quantum mechanics, I am not a committed believer in the relative state interpretation.

The following pages were scanned from a hard copy, and the address in them is obsolete.

—Los Alamos, December 2002 and January 2004

*barnum@lanl.gov

THE MANY-WORLDS INTERPRETATION OF QUANTUM MECHANICS: PSYCHOLOGICAL VERSUS PHYSICAL BASES FOR THE MULTIPLICITY OF "WORLDS".

Howard N. Barnum
6401 Academy NE, #168
Albuquerque, NM 87109
(505) 828-2567

THE MANY-WORLDS INTERPRETATION OF QUANTUM MECHANICS: PSYCHOLOGICAL VERSUS PHYSICAL BASES FOR THE MULTIPLICITY OF "WORLDS".

Abstract. A crucial distinction in discussions of the many-worlds interpretation of quantum mechanics (MWI) is that between versions of the interpretation positing a physical multiplicity of worlds, and those in which the multiplicity is merely psychological, and due to the splitting of consciousness upon interaction with amplified quantum superpositions. It is argued that Everett's original version of the MWI belongs to the latter class, and that most of the criticisms leveled against the MWI, in particular that it is illogical or incoherent, are not valid against such "psychological multiplicity" versions. Attempts to derive the quantum-mechanical probabilities from the many-worlds interpretation are reviewed, and Everett's initial derivation is extended to show that these are the unique possible probabilities. But there remains a challenge for proponents of the MWI: to show that their interpretation requires probabilities, rather than merely nonprobabilistic indeterminacy.

Introduction. While it has enjoyed some popularity, particularly among cosmologists, the relative state or many-worlds interpretation of quantum mechanics (henceforth MWI) appears to have fallen into disfavor among those concerned with the problem of interpreting quantum theory. D'Espagnat (1976) has criticized the theory in his seminal book, and Wheeler (1977), one of the best known proponents, has renounced it. Bell has said that "...the many-universes interpretation is a kind of heuristic, simplified theory, which people have done on the backs of envelopes but haven't really thought through. When you do try to think it through it is *not* coherent." (Bell, interviewed in Davies and Brown (1986)), and his view appears fairly widespread. Even among those who view the MWI as consistent, virtually no-one anymore believes with Everett that it shows that "the Schrödinger equation formalism yields its own interpretation." The major objective of this paper is to show that despite such criticism, the MWI is indeed a coherent and consistent interpretation of quantum mechanics arising naturally from the Schrödinger formalism, by developing a "psychological" version of the MWI in which the multiplicity of "worlds" appears as a multiplicity of perspectives on a single physical reality as seen from the points of view of different conscious subsystems[1] of this physical reality. This physical reality will be described by the usual quantum mechanical apparatus of a Hilbert-space statevector evolving over time according to the Schrödinger equation, with no extra physical structure introduced to account for "branching universes," and will be shown to give rise to the multiplicity of "worlds." It will also be shown to imply that there is only one acceptable probability distribution over them, which coincides with the quantum mechanical one.

---

[1]The term 'subsystem' is here meant in its general sense, not necessarily in its special quantum mechanical sense of one factor of a tensor product of Hilbert spaces.

Sections 1 and 2 provide preliminaries, with section 1 reviewing the reasons for wanting an interpretation, such as the MWI, based solely on the Schrödinger equation and rejecting the projection postulate, and section 2 setting forth a variety of such interpretations, including different possible "many-worlds" interpretations, for comparison with the "psychological" version of the MWI which is the main subject of the paper. Section 3 develops this psychological version in detail, and argues that it coincides with Everett's views, while sections 4 and 5 develop it further by comparing its approach to psychological unity through time and across subspaces of Hilbert space to that of other versions of the MWI--in the process setting to rest the fear, expressed by Bell (1981), that the MWI indulges in "solipsism of the present moment." Section 6 reviews criticisms of the MWI and attempts to answer them in the light of the psychological version. Sections 7 through 9 take up the question of the relationship between the MWI and the probabilities defined by the Born interpretation of the wavefunction. The key points in the discussion are section 8's argument that the Born probabilities constitute the only permissible probability distribution over different "worlds" corresponding to different observed experimental outcomes, and section 9's argument that we cannot yet rule out the possibility of nonprobabilistic indeterminacy. A variety of considerations bearing on the choice between the MWI and other "no-statevector-reduction" interpretations are noted as they arise throughout the paper, and some are touched on again in the concluding section. But the main thrust of the paper is that whatever such considerations may suggest about the choice between the MWI and other interpretations, the only problem with the consistency or coherence of the MWI lies in the need to arbitrarily postulate that the indeterminacy of which branch of the statevector one will end up in is of a probabilistic, rather than nonprobabilistic nature. The lack of an argument for this postulate based in the MWI itself, remains an obstacle to the claim that the MWI provides a consistent, coherent interpretation of quantum mechanics which can generate the quantum mechanical predictions from the Schrodinger equation formalism without any extra structure or postulates.

1. Reduction, Schrödinger Evolution, and the Many-Worlds Interpretation. The MWI can be viewed as an attempt to answer the question: why don't we observe superpositions of macroscopically different states? Schrödinger pointed out, with his cat example, that the Schrödinger equation implied that such states could be prepared. The "orthodox" interpretation's answer to the question of why we don't observe them was to postulate a mode of time evolution distinct from the continuous, unitary linear evolution according to the Schrödinger equation, that applied whenever a measurement was made on a system, throwing it from a superposition into one of the eigenvectors of the measured observable (with probabilities given by the squares of the amplitudes of the components of the statevector expanded in terms of these eigenvectors). This "collapse" or "reduction of the wavepacket" creates several difficulties. Particularly bothersome are: (1) there is no rigorous, physical specification of the conditions under which collapse, rather than Schrödinger evolution, occurs, and (2) gedankenexperiments in which one tests whether collapse has occurred by looking for the interference between the superposed alternatives which, on a view rejecting reduction, would persist after measurement, suggest such interference would be observed.[2] In (2), I am thinking specifically of the reversal of the Stern-Gerlach experiment. (For classic discussions of this, see Wigner (1967) and Bohm (1951); for a modern treatment, see Englert, Scully, and Schwinger (1988) and even permutations thereof (1988), (1989).) While it has been argued

[2]Other difficulties are: (3) the orthodox view seems to require the postulation of a non-quantum-mechanically described outside system "observing" a quantum mechanical system, creating difficulties describing the whole universe in quantum-mechanical terms, and (4) it is probably not possible to describe the "reduction of a statevector" in a Lorentz-invariant way (see Aharonov and Albert (1980, 1981)).

that thermodynamic considerations make even the partial reversal of macroscopic measurements impracticable (Daneri, Loinger, and Prosperi (1962); Peres (1980)), that does not resolve the issue of principle.

What inclines us to accept reduction, rather than continual Schrödinger evolution, is that we don't directly observe macroscopic superpositions. One might make an acceptable theory by making physically precise the conditions under which reduction occurs, either by describing an instant where it occurs, or by slightly modifying the Schrödinger equation so that it remains as it is for microscopic phenomena unobserved by macroscopic apparatus, but shades (continuously?) into statevector reduction as the system interacts with larger and larger apparati capable of correlating its state with more and more macroscopic phenomena. Either of these possibilities are in principle, experimentally distinguishable from pure Schrödinger evolution. But if we can find an interpretation of the quantum mechanical apparatus of state-vector and Schrödinger equation which implies that we won't observe macroscopic superpositions, then we won't have to assume reduction. Both Bohm's (1952) theory (which adds a bit to the quantum mechanical apparatus, but doesn't alter its experimental implications) and the MWI are attempts to do this, in different ways.

2. Different Versions of the MWI and Related Interpretations. Different proponents of the MWI present different versions, and are not always clear about what they have in mind. But clarity about one's interpretation and the assumptions behind it are essential to an adequate assessment of it. I will therefore outline several possible versions of the MWI, and some related interpretations which will help clarify the MWI by contrast and which I regard as its closest competitors. All of them attempt to solve the measurement problem by rejecting statevector reduction. Some of these versions were suggested by the incisive analysis of Bell

(1981) but I will concentrate on a distinction which Bell neglected, that between "physical" and "psychological" interpretations of the multiplicity of worlds.

(A) The first version of the MWI is the simplest, and perhaps the closest to Everett's. The structure of the physical world corresponds just to the statevector in Hilbert space, evolving according to the Schrödinger equation. The multiplicity of "worlds" is really a multiplicity of *consciousnesses*, not a multiplicity of physical systems, and the "branching" that occurs is also an essentially psychological process that occurs when a consciousness observes a superposed system: the conscious system goes into superposition, as the Schrödinger equation implies, but thereby splits into several consciousnesses. There is no multiplication of *physical* worlds, only a complexification of the structure of the statevector of the universe, which has differentiated so as to support, indeed given the psychological conditions for the unity of consciousness to require, several consciousnesses where before there was only one. This interpretation will be my primary focus, and I will expound it at length below. (It might be clearer to call this the Relative State Interpretation, reserving the term 'Many-Worlds' for physical-multiplicity interpretations like B and C below; I refer to both types of interpretations as versions of the MWI, since this has become the established practice.)

(B) Realism about the statevector as above, plus the statement that there is a *physical* multiplicity of worlds. The square modulus of the projection of the statevector on an eigenvector of an observable Z is stipulated to give us the proportion (or relative measure) of worlds in which a measurement of that observable yields the corresponding eigenvalue. (The usual probability assertions of quantum mechanics may then be derived from the assumption that we are equally likely to be in any one of the existing worlds). Deutsch (interviewed in Davies and Brown (1986)) may have in mind something of this sort.

(C) Interpretation B, with the stipulation that the worlds that exist correspond to position eigenvectors; this is essentially Bell's (1981) version of Everett's (1957a,b) MWI. This gets rid of the difficulty with (B), that it introduces "observation" into the fundamental physics again in the choice of an observable whose eigenvectors determine the nature of the physically different worlds.

(D) Interpretation (B), plus a linking of particular worlds at one time with worlds at other times, into *physical* trajectories. Just how to select which worlds get linked into trajectories is a problem seemingly susceptible to many solutions; one desirable criterion might be an "anti-solipsism" requirement: that what is in our memory of "events at time t-1" at time t, correspond to what happened *in this branch* at time t-1. (It would seem there is no way we could observationally distinguish theories which violate this requirement from those that don't, however.) A great variety of specifications of trajectories may be possible; a particularly interesting specification gives rise to trajectories like those in the de Broglie-Bohm pilot wave model.

(E) B, but with the various worlds being merely *possible* worlds, and the real world being chosen with a distribution $\|\psi(\mathbf{z})\|^2$ out of the class of possible worlds, where $\mathbf{z}$ ranges over the eigenvalues of the relevant observable Z. This is another of Bell's suggestions, though it is of course no longer a many-worlds interpretation.

(F) Like (E), but at each moment the actual world is chosen out of the comparison class of possible worlds with distribution $\|\psi(\mathbf{x})\|^2$, where **x** ranges over the eigenvalues of position.

(G) Like (D), but there is a single actual trajectory. For a particular choice of form for the trajectories, interpretation G becomes the de Broglie-Bohm pilot wave theory of Bohm (1952).

E,F, and G are of course not many-worlds theories (I would call them all hidden-variable theories); A through D are.

Do any of these interpretations correspond to the classic expositions of the MWI by Everett, DeWitt and others (De Witt, and Graham 1973)? Despite formalized expositions like Everett's "relative state" formulation, the MWI is commonly thought to be rather unclear. The root of the unclarity is Everett's claim that the MWI is the natural interpretation of the quantum apparatus of statevector and Schrödinger equation, to which it adds nothing. This has seemed to conflict with talk of "branching universes" which appear to be an added structure imposed on the evolving statevector, since after all components of statevectors are not linked through time in the quantum formalism in any kind of tree structure. Moreover, if the additional structure really is a branching tree, the problem of reduction of the wavepacket has not really been avoided, but merely reformulated: the splitting of a universe appears to be basically the same process as reduction, with all the same problems as to when, precisely, it occurs. These difficulties stem from a fundamental ambivalence among critics and some proponents of the MWI about whether the "many worlds" reflect a *physical* multiplicity of universes, or simply a division of consciousnesses within one physical system. I will argue that Everett explicitly endorsed the latter view.

3. Version (A) of the MWI: Psychological Multiplicity Without Physical Multiplicity.

3.1 Psychophysical parallelism. The MWI as presented by Everett is an attempt to interpret the statevector and Schrödinger equation without any extra mathematical formalism (and in particular, without the either the projection postulate or a physical multiplicity of worlds). I read Everett as endorsing version (A) of the MWI. On this reading there is one universe, represented by a statevector, which may however be a superposition of infinitely many macroscopically or microscopically different states. Evolution of this state vector is always according to the Schrödinger equation. To the question "Why don't we observe macroscopic superpositions like Schrödinger's cat?" a simple reply may be given. One system measures (observes, if it is conscious) another when its state becomes correlated with the other's through Schrödinger evolution. That is, letting ψ^A represent the apparatus (or observer) state and ψ^S represent the observed system state, when the state of the total system A + S evolves from $\psi^A_0 \otimes (\Sigma_i \psi^S_i)$ to $\Sigma_i(\psi^A_i \otimes \psi^S_i)$. ($\psi^A_0$ represents the initial, nonregistering state of the apparatus, while the ψ^A_i are distinct states which may be thought of as pointer positions (or states in which the observer sees a definite pointer position) or the like.) In attempting to observe a macroscopic superposition like Schrödinger's cat, my state would become correlated with the cat's, and I would go into a superposition of states "me seeing a dead cat" and "me seeing a live cat." Such a macroscopic superposition does not possess the necessary unity to be a "state of consciousness" (using 'state' in a non-QM sense here). Therefore what happens is my consciousness splits into two consciousnesses, for the components "me seeing a dead cat" and "me seeing a live cat," which are both elements of physical reality, *do* possess the unity required to be states of consciousness. And these components (and their subsequent histories, until their next encounters with amplified quantum phenomena) moreover possess the necessary continuity (of memory-contents, e.g.) with earlier and later states to form the history of a consciousness persisting through time-- with the one bizarre aspect that when I observe an amplified quantum superposition (when I interact with it via a Hamiltonian which correlates components of my state to components of the superposition) my consciousness splits or forks into many branches. (Theoretically, though as a practical matter it would be impossible, someone could perform on me the analog of the reversal of the Stern-Gerlach experiment, and the branches could join again, but in such a process the results of measurement would be erased from my memory(s)--along, perhaps, with everything else that happened between the splitting and rejoining.) These branches are branches of my consciousness, and the branching is not a physical process, but one that occurs because of the conditions for personal identity, or identity of consciousness, which are more matters of philosophy or psychology than physics.

It may seem strange to invoke psychological or philosophical criteria in explaining the results of physical processes, but it is not really. Any physical theory needs some crude account of how the theory links up with our observations, in order to be of interest to us (though of course the theory's meaning is not exhausted by statements about observations). These include crude theories about our eyes, bodies, sense-organs, and how they perceive what they do. Advocates of the MWI could argue that in order to get theory properly related to observation (part of interpreting the theory or mathematical formalism) in quantum mechanics, we simply need a slightly more sophisticated "stylized psychology" which takes into account conditions for the unity of conscious experience. This does not mix psychology into physical theory, so that we need psychological concepts like mind in the statement of the basic physical theory; we simply need to know a bit about the way psychological states *supervene* on physical states in order to get the observational implications of the basic physical theory (which however contains no psychological concepts) right. By contrast, it might be argued that some proponents of versions of the more "orthodox" interpretation of QM, with the reduction postulate, do invoke psychological concepts at the level of basic physical theory: for example Wigner (1967, 1976), who suggests (in part because of his acceptance of the reverse Stern-Gerlach argument against early reduction of the statevector) that reduction must not be a purely physical process, but must rather be an effect of nonphysical mind on matter. *This* sort of thing is worrisome in physical theory (particularly as we have no precise way of deciding when we are in the presence of a "mind" or Cartesian "soul"), but the MWI is not guilty of it.

3.2 Everett as a Subscriber to Version (A)

Everett spends a lot of time on formalized relative state theory, and relatively little on the assumptions about psychophysical parallelism which I have argued are crucial to the success of the MWI. It is these assumptions which give certain "relative states" (states relative to particular states of conscious observers) their interest as potential "worlds" from the point of view of particular consciousnesses. Among supporters of the MWI, Zeh is most explicit about this: "The different components represent two completely decoupled worlds.... As the "other" component cannot be observed any more, it serves only to save the consistency of quantum theory.... This interpretation, corresponding to a "localization of consciousness" not only in space and time, but also in certain Hilbert-space components, has been suggested by Everett..." (Zeh 1970) Everett himself, while stressing the importance of psychophysical parallelism and of representing the observer as a physical system within the theory, does not explicitly discuss the role of conditions for the unity of consciousness. There are strong hints that the view sketched above is implicit in his thought, however:

> We thus arrive at the following picture: Throughout all of a sequence of observation processes there is only one physical system representing the observer, yet there is no single unique *state* of the observer (which follows from the representations of interacting systems). Nevertheless, there is a representation in terms of a *superposition*, each element of which contains a definite observer state and a corresponding system state. Thus with each succeeding observation (or interaction), the observer state "branches" into a number of different states. Each branch represents a different outcome of the measurement and the *corresponding* eigenstate for the object-system state. All branches exist simultaneously in the superposition after any given sequence of observations. (Everett 1957a)

In a footnote, he adds "This total lack of effect of one branch on another also implies that no observer will ever be aware of any "splitting" process.

This passage is somewhat ambiguous, and I suspect different interpretations of it and passages like it (Everett 1957b, p. 62n.) have given rise to many different suppositions about what Everett's MWI really is. One thing that suggests the splitting is a psychological splitting of consciousness, rather than some sort of physical splitting of one world into many, is that Everett clearly says that "the observer state "branches"" rather than, say, "the universe "branches"." That Everett always puts the terms 'split' and 'branch' in quotes suggests the same thing. Also, Everett states that throughout the process, there is only one physical system representing the observer.

4. Psychological Unity Through Time in Various Versions of the MWI. Bell's strongest objection to the MWI is that he believes it implies a "solipsism of the present moment," in that the MWI shows that for an observer in given branch of the statevector at a given time, all of his memory-contents will be (statistically speaking) consistent with the usual quantum-mechanical predictions, but, Bell believes, there is no linking up of these time-slices of consciousnesses into histories: no guarantee that the "I" that has certain experiment-results recorded in its memory actually had those experiences (Bell 1981). Although Everett's scheme implies the existence of *an* "I" at a previous time who had those experiences, Bell thinks there is no physical identification of this "I" with the current "I" who remembers the experiences. I think this problem stems from a misunderstanding of the relative importance of physics and psychology/philosophy in the MWI. The primary aim of the MWI was precisely to show that an interpretation of QM observationally indistinguishable from the orthodox one could be constructed by adding to the Schrödinger equation formalism only some assumptions about psychological or philosophical criteria for identity of consciousness. A minimal way of attempting this gives rise to Bell's time-slices. But if we have qualms about the metaphysical sparseness of the world thus construed, we may add additional suppositions about criteria for identity of consciousness through time, and recover a world of persisting (though branching) individual consciousnesses whose current memory-contents are generally reliable indicators of past experiences. (Moreover, we can get this unification through time whether we have a type (A), (B), or (C) version of the MWI: that is, whether one supposes the multiplicity of worlds at any *given* time is physical or psychological.) Thus if we want to require (1) some degree of physical causal connectedness and (2) some degree of consistency of content, such that memories generally correspond to earlier experiences, as conditions for identity of consciousness through time, we can link up the "right" components of the statevector into consciousnesses with branching histories, which are not solipsistic in the way Bell was worried about. Each component of the universal statevector at time t is causally linked to components at later times (requirement 1), and some of these later components share enough conscious content (requirement 2) that they may be linked into a conscious history. It is a mistake to suppose that *physics* should be expected to supply the extra structure which ensures these trajectories link up. (Though if the physics does add trajectories, as in (D), we will want them to be the right ones for psychological continuity.)

At one point Bell likens Bohm's theory to the MWI plus trajectories (and of course, the singling out of *one* trajectory as the physically real one). He sees a major advantage for Bohm's theory in the specification, in the physics, of a trajectory--a single history for the world, which provides a basis for rejecting solipsism. Presumably the idea is that now I *know* I am identical with the previous person whom physics says existed and had experiences I now remember having, because there aren't all these other "versions of me" in other universes who I might worry that I'm *really, metaphysically* continuous with rather than the one whose experiences I remember. But this problem isn't unique to the MWI (There are other people even in a single branch of *this* universe and how, after all, do I know I wasn't *really, metaphysically* one of them a few seconds ago, instead of the person whose experiences I remember? How do I know I was *really, metaphysically ANYBODY* a few seconds ago, even if there are no other "versions of me in other branches of the universe" to worry about?). I think Bell is here expecting physics to solve puzzles which it need not solve. Finally, it is important that we *can* introduce psychological "trajectories" in version (A), since as I will show in Section 8, this is useful in the derivation of the correct probabilities from version (A), a crucial part of any argument for the observational indistinguishability of version (A) from standard quantum mechanics.

5. Psychological Unity and Multiplicity Across Subspaces of Hilbert Space in Various Versions of the MWI. The considerations of the previous sections suggest two ways of viewing the measurement process, corresponding to different views of the psychophysical relationship between brains in different Hilbert-space components and consciousnesses. In one, there is a single consciousness before the interaction, splitting into several afterwards. In the other, there are the same number of consciousnesses before and after the interaction: one for each component of the Hilbert space decomposition in terms of eigenvectors of the relevant observable. Everett appears to plump for the former, and it indeed version (A) of the MWI appears to provide no basis for a psychological multiplicity of consciousnesses *before* measurement. In the case of type (B) or (C) theories, which have physical mutiplicity but no trajectories, one might imagine either that consciousness unifies across physically distinct worlds in which its state is identical, or that it remains numerically distinct.

6. Critics of the MWI.

6.1 Bell. Bell's views on Everett, as I have suggested above, provide an example of how an excessively physical interpretation of Everett's statements about "branching" and "multiple worlds" can lead to trouble. Bell's view of Everett's theory is summarized in the statement that:

> The Everett (?) theory of this section will simply be the pilot-wave theory without trajectories. Thus instantaneous classical configurations x are supposed to exist, and to be distributed in the comparison class of possible worlds with probability $\|\psi^2\|$. But no pairing of configurations at different times, as would be effected by the existence of trajectories, is supposed. (Bell 1981)

Bell then makes a series of points, labelled (A) through (D), about how this Everett (?) theory seems to differ from what Everett (and/or DeWitt) actually say. I think it is fair to say that Bell supposes a coherent version of the Everett theory *must* differ from what Everett and DeWitt seem to be saying on these four points. But in my view, Bell misinterprets Everett, at least, on these points, precisely because he takes his statements too physically.

Bell's point (A) is that it is unnecessary to multiply worlds. Bell begins by making what I have taken to be the main point of the MWI, that "The psychophysical parallelism is supposed such that our representatives in a given "branch" universe are aware only of what is going on in that branch." But he goes on: "Now it seems to me that this multiplication of universes is extravagant, and serves no real purpose in the theory, and can simply be dropped without repercussions. So I see no reason to insist on this particular difference between the Everett theory and the pilot-wave theory-- where, although the *wave* is never reduced, only *one* set of values of the variables x is realized at any instant." Now this is an interesting point--that one could throw out all worlds but one, and if it were true that one would lose nothing essential, one might want to do it. Everett and DeWitt believe that one does lose something; the existence of all branches is essential to the Everett-DeWitt-Graham demonstrations, discussed by Bell under (C), that one can recover the Born statistical interpretation from the formalism itself. Since Bell thinks this demonstration is unsuccessful, I assume he therefore isn't worried about losing this. From his point of view, *both* the hidden variables interpretations and the MWI have to make an additional assumption introducing the correct probability measure. Bell's Everett(?) (and Bohm) do indeed add such an assumption about the selection of the actual world from an ensemble of possible worlds with probability measure on that ensemble given by $\|\psi^2\|$. Moreover for Bell, who sees both the MWI and de Broglie-Bohm as introducing a structure of "worlds" or trajectories in addition to the statevector, we can remove the other worlds from this extra structure without losing the possibilities of interference present in the statevector; for Everett, the statevector is all there is, and we cannot get rid of these other worlds without destroying these interference possibilities.

Another problem is discussed by Bell under (B):

> it could be said that the classical variables x do not appear in Everett and DeWitt. However it is taken for granted there that meaningful reference can be made to experiments having yielded one result rather than another. So instrument readings, or the numbers on computer output, and things like that,

> are the classical variables of the theory. We have argued already against the appearance of such vague quantities at a fundamental level..... It was for this reason that the hypothesis was made of fundamental variables *x*, from which instrument readings and so on can be constructed.... I suspect Everett and DeWitt wrote as if instrument readings were fundamental only in order to be intelligible to specialists in quantum measurement theory. (Bell 1981).

This is closely related to another problem: that to specify a probability distribution over "worlds" we must choose a preferred operator, so that the components of the statevector along its eigenvectors can be taken as the set of "points" on which the measure is defined. But for Everett, the different "worlds" are not physically defined "configurations" (Bell's Everett(?)) or "trajectories" (de Broglie-Bohm), but are worlds-for-consciousnesses. I stress here the fact of consciousness, not the macroscopic nature of instruments. Macroscopic superpositions exist; we just cannot, in the nature of things, perceive them. Instrument readings are important because they are observed by consciousnesses. Thus while the division into worlds is indeed dependent on vague, nonphysical concepts like consciousness, these concepts are *not* imported into the basic physics or the mathematical formalism, because the grouping of Hilbert-space components into distinct worlds is not a phenomenon at that level (though it could be said to have a basis at that level in the possibility of superposition inherent in the QM formalism). Therefore, there is no need in the basic physical formalism to choose a preferred observable whose eigenvectors determine the nature of the various "worlds."

(C) is the contention that Everett/de Witt do not succeed in deriving the desired probability measure over final configurations from the QM formalism. "Everett has to attach weights to the different branches of his multiple universe, and in the same way [as Bohm] does so in proportion to the norms of the relevant parts of the wave function. Everett and De Witt seem to regard this choice as inevitable. I am unable to see why, although of course it is a perfectly reasonable choice with several nice properties." (Bell 1981). This point is very important, since if we can't get the right probabilities straight from the physical formalism then we will have to add something to it, which will probably require mention of a particular decomposition of the statevector into worlds and explicit specification of a measure over those worlds, thus requiring the extra structure of Bohm and Bell's "classical configuration variables x." The validity of Bell's contention, and of various attempted derivations of the probability measure by MWI proponents, are considered at length in sections 7-9 below.

(D) is the question of trajectories and branching. I have already indicated that branching is not a process represented in the fundamental physics, and Bell is mistaken to attribute to Everett the view that it is represented there. Bell also mentions "the [lack of] association of a particular present with a particular past". I have already discussed this as well: one may suppose such association occurs in the unification of consciousness, rather than by explicitly putting trajectories in terms of "configuration variables" into the physics.

6.2 Wheeler, Wigner, and d'Espagnat. Aside from Bell's, the most influential criticisms of the MWI have probably been those of Wheeler, Wigner, and d'Espagnat. The portions of their criticisms which are relevant to version (A) of the MWI are quite similar. They object to the existence of branches of the statevector in which "we" have experiences different from those we actually have. Thus Wheeler:

> ...great difficulties would seem to arise in giving any well defined meaning to the term 'state of memory of the observer'. Thus as Bohr always emphasized, any attempt to "push the analysis of the mechanism of living organisms [i.e. the consciousness as ultimate observing device] as far as that of atomic phenomena...would doubtless kill [the] animal" [and thus wipe out that very consciousness]. (Wheeler 1977. Brackets Wheeler's).

He then cites Wigner and d'Espagnat. Now it is likely that an attempt to view the consciousness as part of the quantum-system, and, say, to try to reverse a conscious Stern-

Gerlach measurement would, even *theoretically* within the MWI, wipe out that observer's memory of what went on between the measurement and its reversal (though not necessarily the observer-- if one assumes enough finesse to reverse a conscious Stern-Gerlach measurement, one might as well assume enough finesse to keep the observer alive in the process). But this does not invalidate the theory. And if one arranged to have it done on oneself, one might well *know* after reversal, that one had been in a superposition (statevector reduction had not occurred), because of the memory gap. Whereas if reduction had occurred, one would be unable to reverse the measurement and might remain conscious of the measurement result.

Wigner shows some awareness of the psychological basis of the MWI, as he argues:

> The epistemology under discussion postulates that I feel to be in only one of the immensely many states in the linear combination of which I actually am. This means, however, that the other states, and hence the total state vector of the world, are meaningless.
>
> One could postulate, in a similar vein, that the world is homogeneous and isotropic and it is only I who sees differences between different locations and different directions--the real state vector is invariant under all Poincaré transformations. Such ideas appear to me--perhaps wrongly--to be detached from reality. (Wigner 1973).

There is an equivocation on 'I' in the first sentence, between 'I' as whichever consciousness I happen to wind up with after measurement, and 'I' as the *physical* subsystem which gets thrown into superposition after measurement. There is no ground for calling the other states meaningless except an extreme positivism in which only measurement results are real. There are many reasons to reject such a positivism and try to find a realistic interpretation of physical theory, and the justification for introducing the "meaningless" other branches into the MWI is that they enable a realistic (and non-hidden-variables) interpretation of QM. Superpositions are necessary in order to explain interference phenomena; in this regard they are unlike the hypothetical supposition of homogeneity and isotropy of the real statevector to which Wigner compares them in the last paragraph: the latter doesn't help explain anything.

D'Espagnat (1976, Chapter 23) deals with two versions of the MWI, his "solution (a)" and "solution (b)." Solution (a) roughly corresponds to a physical multiplicity view, my (B) or (C), to which d'Espagnat objects because of the need to choose a preferred observable. Solution (b) is one in which "the number of instruments and the number of systems is always conserved." It would therefore appear to correspond to my version (A), but d'Espagnat does not appear to recognize the need for a psychological interpretation of the split.

> Thus we are led to the conclusion that, if the impressions of the observers are just physical properties of these systems--and as such, are described by the state vectors--then these impressions *must* be blurred in the kind of experiments studied here. This conclusion is, however, in obvious contradiction with our immediate experience, since under the conditions of these experiments every one of us knows with certainty that he always has the definite impression of seeing, or not seeing, a signal. (d'Espagnat 1976, p. 273)

The problem with d'Espagnat's conclusion (and the argument that leads to it) is the assumption that there is a single observer consciousness associated with a single physical observer-system and represented in the statevector (the plural in the first sentence refers to a hypothetical ensemble of identical universes (d'Espagnat likes to couch his quantum theory in terms of ensembles), not to the branches of a statevector within a universe). We cannot expect the statevector to be in a "one-to-one correspondence" with *my* single impression; it corresponds to a multiplicity of impressions, which is *not* the blurred impression of a single consciousness, but the distinct impressions of the multiplicity of consciousnesses which my version (A) says must exist after measurement. To expect the objective description of physical reality to contain the "fact" that *I* had *this* particular impression is like expecting it to explain why my consciousness is that of being Howard Barnum, and not of being

Bernard d'Espagnat: in both cases the explanation is roughly that various subsystems of physical reality with a certain structure exist, some sort of crude psychology tells us that each subsystem with such a structure has consciousness, and given that *I* am going to be a consciousness, I have to be one of them and can't be more than one at once, but physics has nothing to say about which one[3], nor should we expect it to.

6.3. Thus it would appear that versions (A), (C) or (D) of the MWI remain competitors for the de Broglie-Bohm hidden variables theory (a variant of G), as do hidden variable interpretation (F) and other variants of (G). There might seem little to go on in making the choice except for metaphysical prejudice (e.g. against unnecessary multiplication of consciousnesses, or against "hidden" variables). But there is also the fact that version (A) does not have to choose a preferred set of configuration variables-- indeed, does not have to add anything to the Schrödinger equation formalism. This may be viewed as an advantage or not, depending on whether one agrees with Bell (1982) that "in physics the only observations we must consider are position observations... It is a great merit of the de Broglie-Bohm picture to force us to consider this fact. If you make axioms, rather than definitions and theorems, about the "measurement" of anything else, then you commit redundancy and risk inconsistency." Speculating, one might even relate this to the possibility (Wigner 1976) that only a few of the mathematically possible Hermitian operators correspond to physically measurable quantities like position and momentum. If this is so, the de Broglie-Bohm measurement theory, and perhaps version (C) of the MWI as well, might explain it, while the standard theory and version (A) of the MWI do not. (Though Bohr's "necessity for using classical variables to describe measurement results" could be considered

[3]But physics combined with the "stylized psychology" of the MWI may enable us to say something about the *probabilities* of being various consciousnesses, when your current one splits due to measurement. See Section 8.

a recognition of it). In addition, there is the fact that we have not, as yet, introduced probabilities into version (A) of the MWI. If they can be derived from this version which does not explicitly mention them, this might be considered an advantage. If they cannot be so derived, we must investigate whether they can be coherently introduced by stipulation (as they are in the physical multiplicity versions of the MWI). The latter question will be investigated in section 7, the former in section 8.

7. The Status Of Probabilities in Version (A) of the MWI. Let's put aside until the next section Everett and DeWitt's contention that the statistical interpretation emerges naturally from the formalism, and consider whether we may simply introduce the square-amplitude measure as a matter of stipulation. Does it make sense to do so? What does it *mean* to assign probabilities to components of the statevector? If one wants to say it has to do with relative frequencies, fine-- but relative frequencies of *what?*

Consider several ways probabilities could be included in a physical theory. (1) Under certain conditions (unambiguously specified in the theory, one hopes) the theory does not predict determinate consequences, but makes a statistical prediction. I.e., the equations of motion may contain transition probabilities. For example, particles make "random jumps." (2) Equations of motion are deterministic, but the initial conditions are probabilistic. This could be described by saying the actual physical world is conceived of as chosen from a (possibly hypothetical) ensemble of possible worlds, with probabilities given by a certain measure. This could be conceived of as (2a)a physically real random process occurring at the "beginning of the world" in which initial conditions are "chosen." Or it could perhaps be due to (2b)our ignorance of initial conditions. In the latter case one wonders if it is proper to call the probabilities part of a genuine physical theory. Perhaps it is, if the physics of the situation explains why we can't know the values and why we should have the subjective probabilities we do. Statistical mechanics is one case in point; Bohm (1953) has argued that his hidden variable theory is another.

The MWI does not fit neatly into any of the above categories. With the MWI, there are determinate initial conditions (initial statevector) and a deterministic law of evolution, so the probabilities must arise from ignorance. But this is not ignorance of an objective physical fact, rather it is ignorance as to which branch of the statevector we are in or will go into--ignorance of a "perspectival fact". In Everett's version of the MWI (my version (A)) the branches "split", rather than being distinct and merely becoming differentiated (section 5), so that it is not a mere matter of finding out which branch one is in, but a matter of finding out which one of the branches one will end up in *after* the split. There is a genuine stochastic event going on; it is, of course, the splitting of one mind into two, as viewed from the perspective of one of the resulting conscious histories. The probabilities are probabilities that "my" consciousness will go with one branch versus the other; and their introduction can be empirically justified by the fact that if we suppose these probabilities obtained in previous, similar, situations where Schrödinger evolution predicts a split (quantum superpositions amplified to affect conscious minds), we get relative frequencies very close to the ones we've actually observed (at least with very high probability, in the limit as the number of observations gets large, which is all any interpretation of quantum mechanics can promise). These probabilities may be thought of as "objective" as opposed to "epistemological" (or as they are sometimes termed, "subjective") probabilities. But they are not the probabilities of objective physical events, but rather of outcomes of a sort of subjective, psychological, perspectival event, the splitting of consciousness.

If the splitting of consciousness were simply a matter of physically distinct but initially similar worlds becoming differentiated, and the probability meant something like the relative numbers of various types of these worlds, the physical basis of the probabilities would be relatively clear. In these interpretations, the worlds in the "comparison set of possible worlds" differ physically, and the probabilities specify a distribution over these different physical possibilities. But in version (A), the probabilities may seem *ad hoc*, necessary for agreement with experiment but a nebulous, hard-to-interpret element in a theory which is, after all, attempting to remedy orthodox quantum mechanics' unclear and/or idealistic view of the world by providing a clear and precise picture of reality.

The difficulty is not *per se* that the probabilities are probabilities of perspectival events rather than probabilities describing events occurring in the world as seen from no particular point of view. It is that we need a coherent picture of how these probabilities of perspectival events (e.g. the probability that "my" consciousness will be the one associated with the "pointer-indicating-z-spin up" component of the measuring apparatus) are grounded in objective physical reality. After all, for other "perspectival facts," one can generally come up with a reasonable explanation of why, given the objective physical situation, things viewed *from here* had to look as they did. One can of course derive the probabilities from a coherent mathematical structure. But we need to justify the identification of these numbers (norms of Hilbert-space components) with probabilities of perspectival events. Since the basic formalism is supposed to be nonprobabilistic, this interpretation of Hilbert-space norms as probabilities really should arise from the basic structure and the nature of the perspectives we can have on it (the notion of splitting consciousnesses), rather than simply being stipulated. Thus, Everett, de Witt, and Graham's claims to have derived the probabilities from the Schrödinger equation formalism become crucial.

8. Deriving the Probabilities From the MWI.

8.1 Everett's derivation. In deriving the square amplitude measure from the formalism Everett's main assumption (1957a,b) is the additivity assumption that if $\psi = \sum\alpha_i\psi_i$ is the component of the wavefunction along the vector ψ, its measure $m(\psi)$ should equal the sum of the measures associated with each of the components in its expansion, $\sum_i m(\alpha_i\psi_i)$. This additivity requirement is justified by a "conservation of probability" argument: we want the probability of being in the "trunk" of a particular tree to equal the sum of the probabilities of winding up on each of its branches. But this argument appeals to trajectories, which are not part of the physical formalism of version (A). One might object therefore that it should be applied to versions of the MWI of type (E) rather than type (A). To this I respond that the "trajectories" are to be interpreted as the psychological trajectories discussed in section 4 above. But what grounds do we have to require the additivity property of these trajectories? Well, we may ask if the apparatus that tells us which branches there are (statevector, Schrödinger equation, and psychological conditions for unity of consciousness) tells us anything about these probabilities. Let a "memory-sequence" $\mathbf{a} = (a_1,a_2,\ldots,a_n)$ be a sequence of observed outcomes of n experiments Note that the sum of probabilities, given that I am in the branch corresponding to a memory-sequence **a**, of going into branches where my memory-sequence is **a** extended by one of the $a_{n+1,k}$ where k indexes the possible outcomes of the $(n+1)$th experiment, which is about to be performed, is one. This is a fact about the way consciousness unifies through time: physical subsystems (in a nontechnical sense) in which one was about to perform a certain experiment unify with later subsystems in which that experiment is complete with a certain outcome, to create a consciousness with a history (or in this case, a consciousness branching into several consciousnesses with histories). If your consciousness starts out in one such system which then becomes correlated with an amplified quantum phenomenon, you can be sure its history will be continued on one of the resulting branches. The reasoning is essentially the same as that which justifies assuming that the continued existence of a physical body and brain with the right relations to one's body and brain at earlier times ensures one's continued conscious existence. And the conclusion is just Everett's additivity assumption.

But Everett's derivation of a measure also makes the seemingly *ad hoc* assumption that the measure m of a given component must be a function of its norm a_i. (He argues that it must be a function of its modulus in fact, since we may simultaneously multiply all weights by an arbitrary modulus-one complex phase factor without changing the physical system represented). But, we might ask, why not let m be a function from $\mathbf{R}^k$ to $\mathbf{R}$, instead of $\mathbf{R}$ to $\mathbf{R}$, a function of *all* the weights in a decomposition, or perhaps even a function of both weights and eigenvectors?

8.2 An improvement on Everett's derivation. To deal with this objection we may keep the additivity assumption and replace the remainder of Everett's derivation with an appeal to Gleason's theorem (Gleason 1957). This theorem states that for each additive measure on the closed subspaces of a Hilbert space $\mathbf{H}$, there is a density operator of the usual quantum-mechanical sort, from which the probability measure can be derived in the usual way). Gleason's theorem does not in itself show that we must assign to a system described by a statevector ψ the probabilities described by the *corresponding* density operator $|\psi\rangle\langle\psi|$ (i.e. the square amplitudes of the expansion of *that* statevector in terms eigenvectors of the measured operator). So that if we had a one-to-one mapping Ω from statevectors to density operators, we might generate a different probability measure, by assigning to a system with statevector ψ the probabilities corresponding to some density operator other than $|\psi\rangle\langle\psi|$, like $\Omega(\psi)$. But a simple argument shows that in the MWI, we cannot do this. It is only reasonable that if a statevector has a zero projection on the subspace spanned by a certain eigenvector, there can be no world corresponding to that subspace in the MWI. (To continue our analogy between spatial-region and Hilbert-space subspace localization of consciousness, if a certain spacetime region is empty, we can hardly assign consciousness to its contents. It seems similarly unreasonable to assign consciousnesses to "empty" subspaces of the Hilbert space.) Therefore we must assign probability zero to worlds corresponding to such empty subspaces. It follows from the axioms of probability that we must assign probability one to a subspace if the statevector lies in that subspace. The only measure satisfying these conditions and additivity is the square-amplitude measure corresponding to the statevector. (Consider the one-to-one function from **H** into the space of density operators mentioned above. Since density operators are Hermitian, they have real eigenvalues and any density operator ρ can be represented as $\rho = \Sigma_i p_i |\psi_i\rangle\langle\psi_i|$, where the ψ_i are an orthonormal basis for **H** and the p_i are between 0 and 1 and sum to 1. Letting the actual statevector be called ψ_0, and letting the $|\psi_k\rangle$ be normalized states orthogonal to $|\psi_0\rangle$ and forming an orthonormal basis together with it, we consider measurement of an observable A which has this basis as its set of eigenvectors. Letting α_k be the eigenvalue corresponding to an eigenvector $|\psi_k\rangle$, we require that $P_A(\alpha_k|\rho)$, the probability that a measurement of A yields α_k if the density matrix is ρ (calculated according to the usual rules for deriving probabilities from density matrices) is 0 for $k \neq 0$, and 1 for $k = 0$. This means that $\Sigma_i p_i \|\langle\psi_k|\psi_i\rangle\|^2 = 0$ for $k \neq 0$. For this to be true, since both the p_i and the $\|\langle\psi_k|\psi_i\rangle\|^2$ are nonnegative, for each i either p_i or $\|\langle\psi_k|\psi_i\rangle\|^2$ must be zero. Since this is true for all i and k, it means that the only $|\psi_i\rangle$ which can enter the representation of the density matrix with nonzero weight p_i are those which are orthogonal to all the $|\psi_k\rangle$ with nonzero k. The only such (normalized) vector is $|\psi_0\rangle$, so the only vector which can enter the representation of the density matrix is $|\psi_0\rangle$, the actual statevector. The only allowable density matrix is thus $|\psi_0\rangle\langle\psi_0|$, the density matrix corresponding to the statevector.) We now have an acceptable alternative derivation of probabilities from the formalism (stressing the psychological nature of the trajectories behind Everett's additivity requirement, and adducing Gleason's theorem and the 0-1 probability constraints imposed by the MWI).

8.3 The De-Witt Graham derivation. Even later proponents of the MWI, de Witt and Graham in particular, felt Everett's arguments for the square-amplitude probabilities to be inadequate, and presented arguments they felt were better. For instance, there is an argument attributed by De Witt (1971) to Graham (and well presented in d'Espagnat (1976)). (Everett states the result, while laying less stress on it than do DeWitt (1971) and Graham (1973)). In this argument, one examines relative frequencies of results in ensembles of identical experiments, and shows that in the limit as the number of experiments goes to infinity, the difference between the observed frequencies and those predicted by standard QM can be made as small as desired, except for a "negligible" set of cases which approaches zero measure. The difficulty is that the argument that this set is negligible rests on the fact that the component of the statevector corresponding to such cases approaches zero norm, and the appropriateness of the norm as a measure of relative probability is precisely what we are trying to establish. D'Espagnat seems not to view this as a problem, describing the extra assumption "that vectors with zero norm correspond to nonexisting branches" as "very natural." This is indeed a natural assumption, but it does not necessarily imply that statevectors with very small norm have very small probability, unless one makes the additional assumption that the relative probability of a branch is a uniformly continuous function of its relative norm. This sounds pretty reasonable, but no argument is given for it. A variant of this argument defines a relative frequency operator on infinite tensor products of Hilbert spaces, i.e on vectors representing possible outcomes of infinite sequences of measurements (and superpositions thereof). Its eigenvalues are shown to be equal to the quantum mechanical probabilities. When the formalism is interpreted, this turns out to be the same argument as that just given, and the same objection arises when one asks why the mathematical idealization of a relative frequency operator on infinite sequences should be relevant to actual measurements on finite sequences.

Graham's objection to Everett is worth quoting at length:

> Everett gives no connection between his measure and the actual operations involved in determining a relative frequency, no way in which the value of his measure can actually influence the reading of, say, a particle counter. Furthermore, it is extremely difficult to see what significance such a measure can have when its implications are completely contradicted by a simple count of the worlds involved, worlds that Everett's own work assures us must all be on the same footing.
>
> (To be sure Everett argues that the measure... is unique. But remember that Gleason has shown that the probabilities defined by the Born interpretation, considered as a measure on a Hilbert space, are themselves unique. Nevertheless this (hopefully) does not deter anyone from inquiring into the connection between those probabilities and experiments that measure relative frequency.) (Graham 1973).

Graham then makes roughly the point I made in the previous paragraph about the DeWitt-Graham-Everett demonstration that relative frequencies converge on the quantum mechanical probabilities: that it relies on the probability measure it is intended to derive. Now Graham's objection that "those probabilities" (in Gleason's theorem) or "[Everett's] measure" are not immediately connected with "experiments that measure relative frequency" requires some subtlety in its interpretation. If the claim is that the square-amplitude probabilities need to be related to relative frequencies, then the DeWitt-Graham-Everett argument does as well as can be expected. *Any* attempt to require an interpretation of probabilities in terms of limits of relative frequencies will run up against the problem that these limits turn out to be plims-- probability limits-- so we are stuck with probabilities as primitives. That is no reason to reject a *particular* use of probability theory such as Everett's.

An alternative interpretation of Graham's objection would be that it is not so much the transition from probability to relative frequency as the transition from measure to probability that is questionable. The formalism allows one to define this square-amplitude measure, but who says it tells us anything about probabilities of events? Maybe these events just happen, and there are no probabilities. Although we *could* assume these particular probabilities (and, given the above argument for the uniqueness of the square-amplitude measure, no others) we are not forced to assume *any* probabilities, so their introduction is "going beyond the formalism," as surely as in versions (C) or (D) or de Broglie-Bohm. As DeWitt (1971) says: "although reality as a whole is completely deterministic, our own little corner of it suffers from indeterminism. The interpretation of the quantum mechanical formalism (and hence the proof of Everett's metatheorem) is complete only when we show that this indeterminism is nevertheless limited by rigorous statistical laws." By examining the way our own little corner fits into the whole, we can apparently derive *what* those statistical laws would have to be, but it is not clear we can derive *that* such laws must apply, unless we are willing to argue that we must assign probabilities when there is a unique acceptable distribution.[4]

---

[4]Graham (1973) has attempted to derive the quantum mechanical probabilities for experiments measuring macroscopic observables by using the quantum statistical-mechanical assumption of equal *a priori* probabilities and random phases for eigenstates. If we are willing to accept that consciousness is necessarily a macroscopic phenomenon, then this argument, if valid, would allow us to base the probabilities on foundations no less secure than those of statistical mechanics. However, the major justification for the Gibbs principle is that the statistical mechanics constructed from it works, empirically. Perhaps the fact that the quantum mechanical probabilities derived from the mere assumption that we must have *some* probabilities also work empirically, makes that assumption equally secure, so that there

9. The Option of Refraining From Assigning Probabilities. The option of claiming that the MWI formalism still gives us no picture of how these probabilities of perspectival events arise, and perhaps even no grounds for assuming such probabilities exist, could correspond to "complete ignorance"or "true uncertainty" in the decision-theoretic sense (Luce and Raiffa (1957, pp. 12, 275-8)), or as I will call it, "nonprobabilistic indeterminacy".

There are several possible arguments against this.

(1) Assuming there are no probabilities makes the theory empirically inapplicable, while the mere assumption that there are *some* probabilities, plus the Schrödinger equation/statevector formalism, yields the empirically *correct* probabilities. If we take this approach, we fail to explain what it is about the underlying reality described by the MWI that makes us choose the probabilistic, rather than the nonprobabilistic, version of the MWI. Furthermore, nonprobabilistic indeterminacy may not be very useful for making predictions, but as I shall argue below, it is *not* empirically inapplicable.

(2) Nonprobabilistic indeterminacy is necessarily "subjective" in the sense of "subjective (i.e. epistemological) probability." It therefore does not apply here, where the probabilities, as argued in section 7, are "objective" probabilities (though of "subjective", i.e. perspectival, events). This is to claim that nonprobabilistic indeterminacy is always a matter of having no clue about the real probability distribution (possibly degenerate), which always exists. The only justification for not assigning a single distribution would be suspension of judgement between a number of possible distributions. But the uncertainty in the many-worlds account of measurement is not due to our ignorance of something now determined. There is no "fact of the matter" before a measurement, about which branch I will go into, any more than there is a "fact of the matter," (except perhaps in some abstruse metaphysical

---

is no need to appeal to statistical mechanics and require consciousness to be macroscopic.

sense) what the outcome of a measurement will be on a statevector-reduction view. The probabilities (or the uncertainty, if we refuse to assign probabilities) are as good as our knowledge could possibly be, theoretically, and in that sense "objective". (See Section 7.) A variant of this argument points out that these events are repeatable, and thus not the kind of "unique event" which it is sometimes argued is the proper context for nonprobabilistic indeterminacy.

I believe this argument is simply wrong. One might cite, for example, the existence of infinite sequences with no limits of relative frequencies to buttress the possibility that "objective" nonprobabilistic indeterminacy could really exist. To counter this, one could argue that such sequences may result (though with probability zero-- which does not mean impossibility) from probabilistic processes. And one might argue that no finite amount of data can convince us we face such an objective nonprobabilistic indeterministic process, rather than a probabilistic one. Nonprobabilistic indeterminacy would then be acceptable only as a description of our ignorance in situations where we must act without adequate data-- never as part of a description of the world which is a candidate for a scientific theory, viewed as potentially standing up to indefinite amounts of further data collection. Since the MWI is viewed as a candidate for a (possibly final) scientific theory, and since the measurements situations which create (perspectival) indeterminacy within it are repeatable, we would have to view these measurements as situations of probabilistic indeterminacy.

The problem with this is that we *can* in theory have perfectly good evidence that we are facing a nonprobabilistic indeterminate process, even on the basis of a finite amount of data. With an infinite amount of data, we might of course have a sequence with no limit of relative frequency, and since this has zero probability on *any* probabilistic hypothesis, we would be justified in assuming a nonprobabilistic process-- which has nothing to say about the probability of the data, of course, but at least doesn't assign it probability zero. But there are also *finite* data sets which justify the inference that it is extremely unlikely they were produced by a probabilistic process. For instance, divide the data into two sets of roughly equal size by some previously chosen, data-independent rule. If (for large amounts of data N) the relative frequencies in the two subsets differ too greatly, we may conclude the data was probably generated by a nonprobabilistic process. For *any* probabilistic process has a very small probability for this sort of discrepancy-- indeed the probability approaches zero for large N.

(3) We can empirically determine relative frequencies, and perhaps this fact is in itself sufficient to justify assuming that there are probabilities. This is not adequate because the task of the MWI is precisely to find a theoretical structure which reproduces the empirical facts, and this includes reproducing the empirical fact that there *are* probabilities. We want a theoretical argument for this from *within* the MWI.

Conclusion. By taking the "multiplicity of worlds" to occur at the level of consciousness, i.e. of a multiplicity of perspectives on a single reality, I have shown that the MWI is a coherent and consistent interpretation of quantum mechanics. Many of the arguments that the MWI is inconsistent were shown to result from the erroneous notion that additional *physical* structure was needed to account for the multiplicity of worlds. The notion, put forth by Bell, for instance, that the quantum mechanical probabilities were only one of many probability distributions which could be imposed on the MWI was also shown to be incorrect when applied to the "psychological" version. But an important, possibly fatal, difficulty was shown to remain: that of vindicating this unique acceptable probability distribution by showing that so far, no adequate argument is apparent which uses the MWI to rule out nonprobabilistic indeterminacy. Those who take a dubious view of the notion of nonprobabilistic indeterminacy (some Bayesians, for instance) may see no difficulty here. I view it as a major difficulty which should be a focus of future work on the MWI.

Even if it is consistent and succeeds in reproducing the experimental facts, there are reasons for disliking the MWI, for instance that it implies the view that our minds occasionally split into many minds, which we will probably never be able to reestablish contact with, in different branches of the "universal wavefunction." While I certainly don't consider it *a priori* impossible that science should lead us to this belief--as proponents of the MWI believe it does--I do consider this belief sufficiently bizarre *a priori* that it is worth investigating the alternatives. However, it may be more scientifically satisfying to adopt the splitting-consciousnesses view along with a physical theory that is rigorous, precise, and economical, than to adopt the "reduction of the statevector" view which is uneconomical in that it postulates two fundamental and radically different modes of change of the statevector, and fuzzy in that it gives no clear physical criteria for when one mode rather than the other obtains. Therefore interpretations which also reject statevector reduction appear the most promising alternatives to the MWI. I argued that the de Broglie-Bohm-Bell type of hidden variable theory, Bell's Everett(?) version of the MWI (C), its hidden-variables analog (F), and perhaps some many-worlds-with trajectories (D) interpretations remain logically acceptable alternatives. Their major differences with version (A) are the need to introduce probabilities explicitly as part of the physics, the need to add something to the Schrödinger formalism, and the related need to choose a preferred observable. Once it is necessary to do this, one might argue, with Bell, that there is little point in retaining the extra worlds; if so, the major alternatives are version (A) of the MWI and the de Broglie-Bohm-Bell pilot wave theory.[5] Future work needs to examine the consequences and justification of the de Broglie-Bohm-Bell introduction of preferred observables.

Despite my defense of the consistency and coherence of the MWI, my own sympathies are more with de Broglie-Bohm-Bell: I find the multiplication of consciousnesses in the MWI unappealing. This might be termed a mere metaphysical preference. But it is possible that the ultimate best choice of interpretation will be determined by its fruitfulness as the basis for new physical theories connecting quantum theory with gravitation, and "metaphysical" preferences between interpretations in part express different views on their fruitfulness in this endeavor. It is perhaps more likely, as Bell often suggests, that some wholly new way of viewing quantum theory will emerge out of such a synthesis. The currently available realistic interpretations of quantum theory may then be viewed as giving us suggestions about the nature of this future perspective on quantum theory, since they give us different pictures of reality which are consistent with current quantum theory and the experimental facts. The psychological version of the MWI suggests the most conservative possibility: that the basic structure of quantum theory may survive intact and without major additions-- at the price of a multiplication of consciousnesses. Hidden variable theories like de Broglie-Bohm suggest that a more radical revision of the quantum formalism, adding additional structure beyond a wavefunction framework, may eventuate.[6] But whatever one's preferences, if one wants an accurate idea of the various kinds of realities that are compatible with quantum theory, it is important to give the psychological version of the

---

[5](F) is vulnerable to Bell's accusations of solipsism of the present moment and seems to have no compensating advantages over any of the other versions. I omit (E) on the grounds that preferred observables must be chosen on arbitrary, "human-centered" grounds. See Bell (1981).

[6]The two types of theory also differ with respect to the nature of "nonlocality" or "nonseparability" they involve, but there is not space to treat this here.

MWI a place among consistent interpretations of that theory, and to understand the true nature of the remaining difficulties with it.

REFERENCES

Aharonov, Y., and Albert, D. (1980) "States and Observables in Relativistic Quantum Field Theories", *Physical Review D 21*: 3316-3324.

Aharonov, Y., and Albert, D. (1981), "Can We Make Sense out of the Measurement Process in Relativistic Quantum Mechanics?", *Physical Review D 24*: 359-370.

Bell, J.S. (1975), "On Wave Packet Reduction in the Coleman-Hepp Model", *Helvetica Physica Acta 48*: 93-8. Reprinted in Bell (1987).

Bell, J.S. (1976), "The Measurement Theory of Everett and de Broglie's Pilot Wave", in *Quantum Mechanics, Determinism, Causality, and Particles*, M. Flato *et al.* (eds.). Dordrecht, Holland: D. Reidel, pp. 11-17. Reprinted in Bell (1987).

Bell, J.S. (1981), "Quantum Mechanics for Cosmologists", in C. Isham, R. Penrose, and D. Sciama (eds.), *Quantum Gravity 2*. Oxford: Clarendon Press, pp. 611-37. Reprinted in Bell (1987).

Bell, J.S. (1982), "On the Impossible Pilot Wave", *Foundations of Physics 12*: 989-99. Reprinted in Bell (1987).

Bell, J.S. (1987), *Speakable and Unspeakable in Quantum Mechanics*. Cambridge University Press, 1987.

Bell, J.S., and Nauenberg, M. (1966), "The Moral Aspect of Quantum Mechanics", in A. De Shalit, H. Feschbach, and L. Van Hove (eds.), *Preludes in Theoretical Physics*. Amsterdam: North Holland, pp. 279-86. Reprinted in Bell (1987).

Bohm, D. (1951), *Quantum Theory*, Chapter 22. Englewood Cliffs NJ: Prentice-Hall. (Dover reprint, Mineola NY, 1989).

Bohm, D. (1952), "A Suggested Interpretation of Quantum Mechanics in Terms of 'Hidden Variables'", Parts I and II, *Physical Review 85*: 160-93. Reprinted in Wheeler and Zurek (1983).

Bohm, D. (1953), "Proof that Probability Density Approaches ψ^2 in Causal Interpretation of Quantum Theory", *Physical Review 89*: 458-466.

Daneri, A., Loinger, D., and Prosperi, G.M. (1962). "Quantum Theory of Measurement and Ergodicity Conditions", *Nuclear Physics 33:* 297-319. Reprinted in Wheeler and Zurek (1983).

Davies, P.C.W., and Brown, J.R. (eds.) (1986) *The Ghost in the Atom.* Cambridge: Cambridge University Press.

d'Espagnat, B. (1976), *The Conceptual Foundations of Quantum Mechanics*. 2nd ed. Reading, MA: Addison-Wesley. (Addison-Wesley reprint 1989).

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

De Witt, B.S. (1971)., "The Many-Universes Interpretation of Quantum Mechanics", in B. d'Espagnat (ed.) *Proceedings of the International School of Physics 'Enrico Fermi' Course IL: Foundations of Quantum Mechanics*. NY: Academic Press. Reprinted in De Witt and Graham (1973) and Wheeler and Zurek (1983).

Everett, H.D. (1957a), "'Relative State' Formulation of Quantum Mechanics", *Reviews of Modern Physics 29*: 454-62. Reprinted in De Witt and Graham (1973) and Wheeler and Zurek (1983).

Everett, H.D. (1957b), *The Theory of the Universal Wave Function*, Ph.D. thesis, Princeton. Reprinted in De Witt and Graham (1973).

Englert, B.-G., Schwinger, J., and Scully, M.O. (1988). "Is Spin Coherence Like Humpty-Dumpty? I", *Foundations of Physics 18*: 1045.

Fine, A. (1973), *British Journal of the Philosophy of Science 24*: 33-34.

Gleason, A.M. (1957), "Measures on the Closed Subspaces of a Hilbert Space", *Journal of Mathematics and Mechanics 6*: 885-893.

Graham, N. (1973), "The Measurement of Relative Frequency", in De Witt and Graham (1973), pp. 229-253.

Luce, R.D. and Raiffa, H. (1957), *Games and Decisions.* New York: Wiley. (Dover reprint 1989).

Peres, A. (1980), "Can We Undo Quantum Measurements", *Physical Review D 22*: 879-83. Reprinted in Wheeler and Zurek (1983).

Scully, M.O., Englert, B.-G., and Schwinger, J. (1989), "Is Spin Coherence Like Humpty-Dumpty? III. Effects of Observation and Correlation", preprint.

Schwinger, J., Scully, M.O., and Englert, B.-G. (1988), "Is Spin Coherence Like Humpty-Dumpty? II", *Zeitschrift fur Physik D 10*: 135.

Wheeler, J. (1977), "Include the Observer in the Wave Function?", in J. Leite Lopes and M. Paty (eds.), *Quantum Mechanics a Half Century Later*, Dordrecht, Holland: D. Reidel, pp. 1-18.

Wigner, E.P. (1961), "Remarks on the Mind-Body Problem", in I.J. Good (ed.), *The Scientist Speculates*, New York: Basic Books. Reprinted in E. Wigner, (1967), *Symmetries and Reflections*, Bloomington: Indiana University Press, and in Wheeler and Zurek (1983).

Wigner, E.P. (1963), "The Problem of Measurement", *American Journal of Physics 31*: 6-15. Reprinted in E. Wigner (1967), *Symmetries and Reflections.* Bloomington: Indiana University Press, and in Wheeler and Zurek (1983).

Wigner, E.P. (1973), "Epistemological Perspective on Quantum Theory", in C.A. Hooker (ed.), *Contemporary Research in the Foundations and Philosophy of Quantum Theory.* Dordrecht, Holland: D. Reidel, pp. 369-385.

Wigner, E.P. (1976), "Interpretation of Quantum Mechanics", mimeographed notes, revised and reprinted in Wheeler and Zurek (1983), pp. 260-314.

Zeh, H.D. (1970), "On the Interpretation of Measurement in Quantum Theory", *Foundations of Physics 1*: 69-76. Reprinted in Wheeler and Zurek (1983).