Quantum Solipsism and Non-Locality

Travis Norsen
tnorsen@smith.edu
(Dated: November 27, 2014)

J.S. Bell's remarkable 1964 theorem showed that any theory sharing the empirical predictions of orthodox quantum mechanics would have to exhibit a surprising – and, from the point of view of relativity theory, very troubling – kind of non-locality. Unfortunately, even still on this 50th anniversary, many commentators and textbook authors continue to misrepresent Bell's theorem. In particular, one continues to hear the claim that Bell's result leaves open the option of concluding *either* non-locality *or* the failure of some un-orthodox "hidden variable" (or "determinism" or "realism") premise. This mistaken claim is often based on a failure to appreciate the role of the earlier 1935 argument of Einstein, Podolsky, and Rosen in Bell's reasoning. After briefly reviewing this situation, I turn to two alternative versions of quantum theory – the "many worlds" theory of Everett and the Quantum Bayesian interpretation of Fuchs, Schack, Caves, and Mermin – which purport to provide actual counterexamples to Bell's claim that non-locality is required to account for the empirically-verified quantum predictions. After analyzing each theory's grounds for claiming to explain the EPR-Bell correlations locally, however, one can see that (despite a number of fundamental differences) the two theories share a common for-all-practical-purposes (FAPP) solipsistic character. This dramatically undermines such theories' claims to provide a local explanation of the correlations and thus, by concretizing the ridiculous philosophical lengths to which one must go to elude Bell's own conclusion, reinforces the assertion that non-locality really is required to coherently explain the empirical data.

I. INTRODUCTION

Fifty years ago, in 1964, John Stewart Bell first proved the theorem which has become widely known as "Bell's Theorem" [1] but which Bell himself instead referred to as the "locality inequality theorem" [2]. In Bell's own view, the theorem showed that the empirical predictions of *local* theories will be constrained by Bell's inequality (or as Bell himself preferred to call it, the "locality inequality"). Hence, *non-locality* is a necessary feature of any theory which shares the empirical predictions of standard quantum mechanics. In recent decades, the relevant inequality-violating quantum mechanical predictions have been confirmed in a series of increasingly accurate and convincing experiments. [3] It is thus known with reasonable certainty that non-locality is a real feature of the world.

This summary of the situation, however, remains curiously and frustratingly controversial, despite the five decades that physicists and philosophers have had to contemplate and understand Bell's arguments. There remain, for example, many commentators (including, undoubtedly, some in this very volume) who regard Bell's theorem not as a proof of non-locality, but instead as a refutation of determinism and/or the so-called "hidden variables" program and/or some (usually ill-defined) notion of "realism". In general, that is, there remain many commentators (including, particularly troublingly, textbook authors) who assert or imply (often without even realizing that they are flatly contradicting Bell's own understanding) that Bell was simply wrong to claim that non-locality was required by the (now well-confirmed) quantum mechanical predictions. [4–7]

A systematic presentation of Bell's arguments, including some polemics against these (and other) persistent misunderstandings, can be found in Refs. [8, 9]. In the present paper, my goal is to focus on one particular thread of such disagreement with Bell, which has been especially influential in the last decade or so. In particular, I will explore the so-called Quantum Bayesian ("QBist") and Everettian ("many worlds") approaches to quantum theory, both of which purport to provide counter-examples to the claim that non-locality is required to account for the empirical data. In particular I will develop the thesis that, although QBism and Everettism are usually thought of as almost polar opposites (with the latter being one of the more popular *realist* versions of quantum theory, and the former often being considered as rather solipsistic), the two theories' grounds for claiming locality are in fact basically similar and, once brought out into the light, deeply unconvincing. Understanding how and why will then lead to a deeper appreciation of Bell's work.

II. BELL'S ARGUMENT

But before jumping in to a polemical discussion of QBism and Everettism, it will be helpful to briefly review Bell's arguments.

In Bell's original 1964 paper, his main analysis begins where the 1935 argument of Einstein, Podolsky, and Rosen (EPR) [10] had left off. That is, Bell begins by recalling EPR's demonstration that (in Bell's words) "quantum mechanics could not be a complete theory but should be supplemented by additional variables [which would] restore to the theory causality and locality." [1] As Bell goes on to elaborate the argument:

> "[W]e make the hypothesis ... that if the two measurements are made at places remote from one another the orientation of one magnet does not influence the result obtained with the other. Since we can predict in advance the result of measuring any chosen component of [the polarization of particle 2], by previously measuring the same component of [the polarization of particle 1], it follows that the result of any such measurement must actually be predetermined." [1]

I will review the actual EPR argument shortly; for now I just want to stress that Bell's 1964 paper begins by recapitulating the 1935 EPR argument, which Bell takes to have established that a deterministic hidden variable theory was one's *only hope*, if one wanted to explain the predicted quantum correlations in a local way.

In the body of his 1964 paper, Bell then shows that this kind of deterministic hidden variable theory's predictions for a wider class of possible experimental measurements (in which the outcomes are correlated, but imperfectly so) are necessarily constrained by a Bell (or, as Bell called it, a locality) inequality. The mathematical details of this demonstration are well understood so I won't bother to rehearse them here. The main point is just that, contrary to the impression of people who miss the role of the EPR argument in Bell's overall thesis, Bell's 1964 result already establishes the inevitability of non-locality. It does *not* leave open some kind of choice between abandoning locality and abandoning "realism" or "hidden variables" or "determinism", because the overall argument does not *begin* with realistic / deterministic hidden variables. Instead it begins with locality alone, proceeds (via the EPR argument) to establish the necessity of deterministic hidden variables in order to locally explain just the subset of the quantum predictions considered by EPR, and then finally closes off that apparent possibility by showing that this kind of theory cannot reproduce the full slate of quantum predictions.

Looking back, and considering the widespread and persistent confusion about its role in his theorem, it is somewhat unforunate that the two or three sentences I quoted above constitute basically the entirety of Bell's recapitulation of the EPR argument.[1] And of course that argument had been originally made in a not terribly rigorous way (made worse by the fact that the actual EPR paper, written by Podolsky and not seen by Einstein until after its publication, obscured the main argument in Einstein's opinion [11, 12]). But still it is very easy to see that the EPR argument that Bell rehearses is entirely valid and provides the necessary foundation for Bell's subsequent demonstration. Consider for example the case of two spatially-separated and polarization-entangled particles. Suppose in particular that the joint polarization state – for example,

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left[\,|HH\rangle - |VV\rangle\,\right] \tag{1}$$

– is such that, if the polarization of both particles is measured along the same direction, the results are perfectly correlated: either both outcomes are "H" or both outcomes are "V".

The EPR argument can then be understood to proceed as follows. Suppose a measurement is made on the nearby particle, yielding some outcome, say "H". It is then certain that the distant particle, if measured along the same direction, will also yield the outcome "H". We could say that now, after the measurement on the nearby particle, it is clear that the distant particle somehow encodes this outcome, "H", in its internal structure.[2] There are then two possibilities. Either the distant particle possessed this internal structure all along (i.e., even before our measurement on the nearby particle), or it didn't (i.e., it only acquired it after, and evidently as a result of, our nearby measurement). The latter option implies *non-locality*: our nearby measurement caused the distant particle to change its physical state, to acquire a definite value ("H") for a property that was previously somehow indefinite (or just different). The former option, on the other hand, implies the existence of an outcome-determining structure in the distant particle, about which quantum theory is silent. The former option thus implies the *incompleteness* of the quantum mechanical description. EPR simply took locality for granted (no doubt on the basis of Einstein's relativity theory, which is usually taken to prohibit faster-than-light causal influences), so the established dilemma between non-locality and incompleteness immediately suggested that quantum mechanics did not provide a complete description of microscopic reality. This was thus EPR's main conclusion.

For Bell's purposes, though, that is not really the crucial point. The important thing was instead that the *only* way to explain the perfect correlations locally is to attribute outcome-determining properties to the individual particles. These properties, evidently, would vary randomly from one particle pair to the next, but would be fixed (in an appropriately correlated way) once and for all at the source for a given particle pair. (For example, perhaps the first pair is of the type "particle 1 is 'V' and particle 2 is 'V'," then the second pair is of the type "particle 1 is 'H' and particle 2 is 'H'," and so on.) It is easy to see that such deterministic hidden variables are the

[1] That is, it would have been nice if, already in 1964, Bell had anticipated the need for a much sharper and more detailed presentation of the EPR part of the argument (including especially a more generalized and more precise formulation of the crucial "locality" premise, along the lines of what he would indeed give later, in 1976 and 1990).

[2] For example, in ordinary quantum mechanics, the quantum state of the distant particle will be $|H\rangle$.

only way to account locally for the perfect correlations: any residual indefiniteness in either particle would either at least sometimes spoil the perfect correlations (if the indefiniteness was resolved locally during the subsequent measurement procedure) or would involve some physical process in which the state of one particle was nonlocally affected by the distant measurement process or result. So we have to choose between non-locality and local deterministic hidden variables. Or equivalently, the only way to avoid non-locality (in the face of the EPR correlations) is to embrace local deterministic hidden variables.

In any case, it should be easy to see – even without a formal definition of locality and a formally rigorous version of the argument – that the EPR argument is entirely valid and that, indeed, as of Bell's writing in 1964, it was already established that the only hope for a *local* explanation of the quantum correlations was a deterministic local hidden variable theory... and that, therefore, by showing that such a theory could not reproduce the quantum predictions in more general situations, Bell had proved that no local explanation of the quantum correlations was possible, full stop. Or as he put it already in the introduction of his 1964 paper: "It is the requirement of locality ... that creates the essential difficulty." [1]

As a matter of social/historical fact, however, none of this was clear. The vast majority of commentators simply missed, or misunderstood, the role of the EPR argument in Bell's reasoning and/or wrongly believed the EPR part of the argument to be invalid. Bell thus exerted a great deal of effort in the subsequent decades to clarify his reasoning and to make all of the relevant assumptions more rigorous and explicit. This is not the place for a systematic presentation of his subsequent clarificatory efforts, but I will mention Bell's classic 1981 "Bertlmann's Socks and the Nature of Reality" in which he remarks, in a footnote:

> "My own first paper on this subject [i.e., the 1964 paper] starts with a summary of the EPR argument *from locality to* deterministic hidden variables. But the commentators have almost universally reported that it begins with deterministic hidden variables." [13, emphasis in original]

The same paper also includes the following admirably clear recapitulation of the EPR argument

> "It is important to note that to the limited degree to which *determinism* plays a role in the EPR argument, it is not assumed but *inferred*. What is held sacred is the principle of 'local causality' – or 'no action at a distance'. Of course, mere *correlation* between distant events does not by itself imply action at a distance, but only correlation between the signals reaching the two places. These signals, in the idealized example of Bohm [involving the perfect polarization correlations], must be sufficient to *determine* whether the particles go up or down. For any residual undeterminism could only spoil the perfect correlations." [13]

as well as this beautiful presentation of the overall argument for non-locality:

> "Let us summarize once again the logic that leads to the impasse. The [EPR-Bohm] correlations are such that the result of the experiment on one side immediately foretells that on the other, whenever the analyzers happen to be parallel. If we do not accept the intervention on one side as a causal influence on the other, we seem obliged to admit that the results on both sides are determined in advance anyway, independently of the intervention on the other side, by signals from the source and by the local magnet setting. But this has implications for non-parallel settings which conflict with those of quantum mechanics. So we *cannot* dismiss intervention on one side as a causal influence on the other." [13]

In addition to thus clarifying the reasoning he had used in the 1964 paper, Bell's other main clarificatory innovation (in the decades after the theorem was first proved) was a careful and explicit and formal definition of "locality" (i.e., "local causality").[3] This allowed more formal and explicit demonstrations of both the EPR argument (from locality to deterministic hidden variables) and Bell's "locality inequality theorem". The next section briefly recalls Bell's formulation of "locality" as it will play an important role in our subsequent discussion of QBism and Everettism.

III. BELL'S FORMULATION OF LOCALITY

Bell's first attempt at an explicit space-time formulation of the principle of locality (i.e., "local causality") occurs in his 1976 "The theory of local beables" [15]. The formulation is in terms of his neologism, "beable", which Bell introduced as a foil to the "observables" that play a central role in the mathematical formulation of ordinary quantum theory. As Bell points out, "It is not easy to identify precisely which physical processes are to

[3] It is sometimes suggested that, for Bell, "locality" and "local causality" are logically and/or conceptually distinct notions, and that sensitivity to this distinction provides a rational basis for the types of views I have been here criticizing as mistakes/confusions. See for example Ref. [14]. In my opinion, though, there is overwhelming evidence that Bell used "locality" and "local causality" interchangeably and was only ever interested in the single, unitary concept that is roughly captured by the idea of "no faster-than-light causal influences". The Preface that Bell wrote for the (first edition) publication of *Speakable and Unspeakable* is particularly revealing on this point. [2]

be given the status of 'observations' and which are to be relegated to the limbo between one observation and another." The beables of a candidate theory are those elements which are supposed to correspond directly to physically real structures posited to exist (independent of any "observation") by the theory. Indeed: "'Observables' must be *made*, somehow, out of beables." [15]

As an example, Bell cites the electric and magnetic fields in classical electromagnetic theory:

> "In Maxwell's electromagnetic theory, for example, the fields **E** and **H** are 'physical' (beables, we will say) but the potentials **A** and ϕ are 'non-physical'. Because of gauge invariance the same physical situation can be described by very different potentials. It does not matter that in Coulomb gauge the scalar potential propagates with infinite velocity. It is not really supposed to *be* there." [15]

The fields **E** and **H** are examples of *local* beables in the sense that they are localized in delimited space-time regions (points, actually). Such local beables are to be contrasted with non-local beables – objects that a candidate theory posits as physically real, but which are not localized in space-time. (The wave function of a many-particle quantum system – if it is considered a beable in some version of the theory – would be an example of a non-local beable: it is a function on the 3N-dimensional configuration space of the N particle system, not in any sense a field that realizes definite values at points in 3+1-dimensional space-time.)

The above passage also illustrates one of the crucial (and neglected) points underlying Bell's formulation: since "locality" and "non-locality" refer to physical processes posited by candidate theories (and in particular the issue of whether causal influences propagate always at or slower than the speed of light, or instead in some cases exceed that speed), it is simply hopeless to try to diagnose whether a theory is "local" or "non-local" until it is made crystal clear what the theory says *exists* in ordinary 3D physical space. The *ontology* (i.e., the *beables*) of the theory, that is, must be clearly and explicitly articulated before one can possibly make any judgment about its status vis-a-vis locality. As Bell put this point in 1976:

> "It is in terms of local beables that we can hope to formulate some notion of local causality." [15]

He might have written (and I think he certainly believed) that it is *only* in terms of local beables that the idea of locality can be clearly formulated. Or, as he put this point in the 1987 preface to the 1st edition of his collected papers:

> "If local causality in some theory is to be examined, then one must decide which of the many mathematical entities that appear are supposed to be real, and really here rather than there." [2]

It will be crucial to appreciate that "here" and "there" refer to locations in ordinary three-dimensional physical space. Bell's point is that positing a clearly-articulated ontology of *local beables* is a pre-requisite for any discussion of a theory's status vis-a-vis *dynamical locality*.

In the 1976 paper, Bell first discusses "local determinism" – the idea that a complete specification of local beables in a space-time region that closes off the past light cone of some region will uniquely determine the beables in that region. (Maxwell's electromagnetism, for example, has this property.) Bell then specifically introduces "local causality" as a more general notion intended to capture the absence of faster-than-light causal influences for any theory, whether deterministic or not. For reasons (partially?) alluded to in a footnote of his 1977 remarks on "Free variables and local causality" Bell subsequently changed, in a subtle way, his formulation of local causality. [16] The updated formulation first appeared in a footnote of his 1986 "EPR correlations and EPW distributions" [17] and then received a much more careful and elaborate treatment in his 1990 "La nouvelle cuisine", the last paper to appear before his tragic and untimely death. [18]

Here is Bell's brief 1986 formulation:

> "In a locally-causal theory, probabilities attached to values of local beables in one space-time region, when values are specified for *all* local beables in a second space-time region fully obstructing the backward light cone of the first, are unaltered by specification of values of local beables in a third region with spacelike separation from the first two." [17]

In the 1990 paper, Bell illustrates the notion with the figure I have reproduced here (with Bell's caption) as Figure 1. He then formulates the principle of local causality as follows:

> "A theory will be said to be locally causal if the probabilities attached to values of local beables in a space-time region 1 are unaltered by specification of values of local beables in a space-like separated region 2, when what happens in the backward light cone of 1 is already sufficiently specified, for example by a full specification of local beables in a space-time region 3...." [18]

It is perhaps also worth quoting Bell's subsequent clarificatory notes:

> "It is important that region 3 completely shields off from 1 the overlap of the backward light cones of 1 and 2. And it is important that events in 3 be specified completely. Otherwise the traces in region 2 of

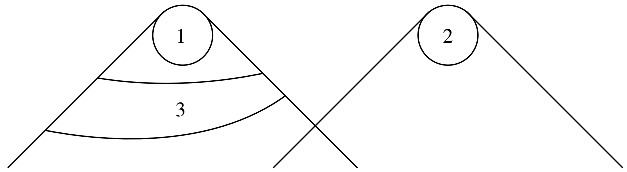

FIG. 1: "Full specification of what happens in 3 makes events in 2 irrelevant for predictions about 1 in a locally causal theory." [18]

> causes of events in 1 could well supplement whatever else was being used for calculating probabilities about 1. The hypothesis is that any such information about 2 becomes redundant when 3 is specified completely." [18]

Any reader not already familiar with this is urged to read Bell's 1976 and 1990 papers (as well as my own detailed analysis in Ref. [9]) to appreciate the systematic clarity of Bell's formulation.

The logical and pedagogical benefits of this explicit formulation of locality are enormous. It is straightforward, for example, to rehearse a fully rigorous version of "the EPR argument from locality to deterministic hidden variables" [13] and thus solidify the foundation of Bell's original 1964 theorem. Alternatively, the explicit formulation of locality also allows a rigorous derivation (which does not rely on perfect anti-correlations and does not introduce deterministic hidden variables as a middle step) of the so-called CHSH inequality. (See Refs [8, 9, 18] for details.) All told, then, Bell's work to explicitly articulate the relativistic notion of "no superluminal action at a distance" makes an essentially airtight case for his conclusion that non-locality is required by the quantum mechanical predictions (and, more importantly, we now know, by actual experiment).

It is somewhat surprising, therefore, that so many people still deny that Bell proved the necessity of nonlocality. As I have suggested above, this is largely a result of plain ignorance about what Bell actually did. [19] Hopefully the above summary and the growing body of good literature on the subject will help improve the situation. But here I want to turn to focus on a rather different category of disagreement with Bell's claim to have demonstrated the empirical necessity of non-locality – disagreement, that is, which is based not so much on a simple failure to understand the logical structure of Bell's arguments, but based instead on an implicit rejection of Bell's point that a clearly-articulated ontology of local beables in three-dimensional physical space is a *prerequisite* to any meaningful analysis of a theory's status vis-a-vis locality.

IV. SOLIPSISM AND FAPP SOLIPSISM

Let us then turn to analyzing several quantum worldviews that, in one way or another, fail to meet Bell's prerequisite. We'll begin by establishing a simple, if somewhat bizarre, point of principle. It is possible to elude Bell's conclusion (that non-locality is required to explain the observed correlations) by adopting the philosophical viewpoint known as "solipsism" according to which nothing outside of one's own subjective conscious experience actually exists. Although it would be hard to name any actual person (philosopher or otherwise) who endorses solipsism fully, the idea can be regarded as a kind of fully consistent implementation of Berkeley's "esse ist percipi" and Descartes' combination of radical doubt and the supposed prior certainty of consciousness. The idea is to regard sense experience – which is normally regarded as experience *of* a material world that exists independent of any conscious awareness – as being instead like a hallucination with no external object at all.

I do not want here to engage with the extensive philosophical literature on the reasonableness and/or refutability of solipsism. I think the situation was pretty well summed up by Bell's remark that

> "Solipsism cannot be refuted. But if such a theory were taken seriously it would hardly be possible to take anything else seriously." [20]

Instead I simply want to note that if one (for whatever reason) absolutely refused to allow that non-locality might really be a feature of the external physical world, adopting solipsism would provide *a* logically possible basis for that stance: if there simply *is* no external physical world, then clearly Bell cannot have established that there are real faster-than-light causal influences in it.

The problem with this stance, of course, is that it commits one to denying quite a lot more than just the alleged non-local causal influences that, in the scenario I've just described, motivate its adoption. The solipsist also denies the existence of tables, trees, planets, other solipsists, and even his own physical body. Within the realm of physics, the solipsist not only denies the physical reality of the measurement outcomes which violate Bell's locality inequality and thus allegedly provide the evidence for Bell's conclusion of faster-than-light causal influences; the solipsist also denies, for example, the real existence of the facts we usually interpret as evidence for the existence of (light-speed) electromagnetic causal influences, the real existence of the facts we usually interpret as evidence for the existence of (sound-speed) verbal influences, etc. He denies, in short, the existence of *everything*. All of it, the whole apparently real physical world that constitutes the (usually-assumed) subject matter for the science of physics, is instead a mere delusion or fantasy or hallucination for the solipsist. Undoubtedly this is what Bell had in mind in remarking that solipsism makes it impossible "to take anything else seriously".

A key point about the solipsist's position here is that although he eludes the need to admit the existence of Bell's non-local causal influences, he does not in any sense retain a local description of the world. He retains no description of the world at all – so the distinction between "local" and "non-local" accounts of the world simply does not apply. That is, although the solipsist's stance is not non-local, it is not, thereby, in any meaningful sense, local. It is neither local nor non-local. In denying the very existence of an external physical world, the solipsist removes any possible meaning from the question of whether causal influences in the external physical world propagate sometimes (the non-local case) or never (the local case) faster-than-light.

The entire reason we are having any of this discussion is of course that non-locality is very difficult to reconcile with a fully relativistic account of physical goings-on in three dimensional space and time. Someone who felt that the evidence in favor of relativity was very strong might thus be willing to trade something – even something quite significant – for a way to elude Bell's conclusion that non-locality is real. But solipsism, as a response to Bell, seems completely crazy: the solipsist trades *everything*, including whatever possible basis he might have had for believing in the relativistic account of physical goings-on in space-time, i.e., whatever possible basis he might have had for wanting to avoid non-locality in the first place. [22] Presumably that explains why, as I said, nobody really endorses solipsism, as a response to Bell's arguments or otherwise. Nevertheless, in principle if absurdly, one could respond to Bell's claims by adopting solipsism as a way to elude nonlocality.

Let us then contrast literal solipsism with a category of views that I will call "FAPP Solipsism". "FAPP" here, following Bell, means "For All Practical Purposes". [21] The most widely-known version of FAPP Solipsism is the notorious brain-in-a-vat scenario, memorably dramatized in the movie "The Matrix", in which it is supposed that you might be deluded about almost everything you believe because, instead of arising from causal contact with a real external world, your conscious sensations might instead result from electrical signals fed to your brain (which is kept biologically viable in a vat of appropriate fluid) by an evil scientist with a supercomputer. Such a view is clearly not literally solipsist since its very formulation presupposes the existence of physically real brains (from which subjective conscious experiences somehow arise), vats, biologically auspicious fluids, electrical signals, evil scientists, and/or supercomputers.

And yet the brain-in-a-vat scenario just as clearly has much in common with literal solipsism. Just as with adopting solipsism, considering that one might really be a brain-in-a-vat is a way to elude not just some particular undesirable conclusion (such as non-locality) but anything one happens to want to elude. For example, if, for whatever reason, you are not only bothered by faster-than-light causation (ala Bell) but also bothered by light-speed causation (ala Maxwell), it is easy enough to assert that that, too, is merely a delusion implanted in your unsuspecting brain by the evil scientists. Or if, for some reason, you have always been unable to accept that the Earth goes around the Sun (ala Copernicus) rather than vice versa, it is easy enough to concoct a story in which the evil scientists live (and set up their diabolical laboratory) on the surface of a planet which rests comfortably at the exact center of a series of concentric, rotating, crystalline spheres.

The point here is that as soon as you allow the possibility that your ordinary perceptual experience might not really be *of* an external physical reality, but might instead be a hallucination fed into your brain by an evil scientist, literally every aspect of the conjectured real world (where the evil scientists are supposed to live and work) becomes purely arbitrary. You can literally make up whatever you want, because the usual epistemological burden of providing empirical (i.e., ultimately, perceptual) evidence for the various aspects of one's proposed world-picture has been short-circuited by the assumption that any perception-based claim is in fact a hallucination (or is, at any rate, otherwise delusory). Note, for example, that the brain-in-a-vat theorist really has (by his own implied epistemological standards) no grounds whatever for believing in the real existence of brains, including his own. After all, whatever evidence you take yourself to have for believing that there exist real human beings, with internal organs including brains, out of whose complex electro-physiological structure conscious experience somehow emerges, is – if you are a brain-in-a-vat theorist – just another set of hallucinations fed to you by the evil scientists.

The brain-in-a-vat theorist is not exactly a solipsist, because he claims to believe that there is a physical reality out there independent of our conscious experiences. But *he might as well be a solipsist*. He claims to believe in a real physical world (with evil scientists and vats and brains in it), but he can, by his own standards, have no grounds whatsoever for those *particular* beliefs about the real world. That is, his beliefs about the physical world (including, I think it must be admitted, even the belief that conscious experience can only emerge, somehow, from an appropriately brain-ish sort of physical object) are entirely arbitrary. They could all be changed, or dropped entirely, without cognitive consequence. The brain-in-a-vat theorist may thus be descrbed as a "FAPP solipsist".

We might reformulate the similarity (between literal and FAPP solipsism) as follows: insofar as they purport to account for our experiences (and thereby achieve a kind of empirical adequacy), literal solipsism and FAPP solipsism both engage directly with our subjective conscious experiences, rather than with the external material facts that those experiences are normally regarded as experiences *of*. They both, that is, purport to explain our subjective conscious experiences – that are "as if of" certain external facts – while denying the real existence of those particular external facts. The point is that whether they then posit some *other* external facts (FAPP solipsism) or not (literal solipsism) is, FAPP, irrelevant, since any other posited facts are necessarily completely arbitrary anyway and can hence play no genuine role in justifying the claimed account of conscious experience.

Considered as a possible response to Bell's claim to have established the real existence of non-local (faster-than-light) causal influences, the essential point is that FAPP solipsism does not attempt to provide any (local) physical explanation for a certain pattern of physically real measurement outcomes (instantiated, say, as the positions, at different times, of some physically real instrument pointers) which violate Bell's inequality. Instead, the FAPP solipsist only needs to provide a story according to which a subjective conscious experience – that is as-if-of such measurement outcomes – would arise. The FAPP solipsist's attempt at achieving empirical adequacy (i.e., consistency with experience) occurs, that is, inside of consciousness – at the level of subjective conscious experiences – rather than in the physical world itself. The real physical world, for the FAPP solipsist, may be nothing at all like the one we ordinarily believe in on the basis of ordinary perception: it need not include Bell's non-local causal influences, it need not include measuring instruments with pointers at all, it need not even be a world with three spatial dimensions. Any kind of world at all will do just fine, so long as one includes, as part of the story, the emergence of subjective conscious experiences which fool the inhabitants of the world into thinking they live in a three-dimensional world populated by cats, trees, planets, and Bell-inequality-violating pointer positions.

I have said that nobody is really a solipsist. I also don't know of anybody who has attempted to refute Bell's reasoning by claiming that we are all really just brains-in-vats. So the immediate conclusion here is a purely hypothetical one: if somebody were to attempt to refute Bell's arguments on these grounds, we wouldn't (and shouldn't) take them very seriously. Bell demonstrated that if you regard the familiar three-dimensional world of everyday perception as physically real, and if in particular you "include the settings of switches and knobs on experimental equipment, the current in coils, and the readings of instruments" as among the *local beables* of your theory – as, he thought, any serious and empirically viable theory *must* do [15]) – then your theory will have to involve a specific sort of non-locality in order to achieve empirical adequacy. Responding to that demonstration by saying "Aha, but maybe there isn't any physical reality at all!" hardly constitutes a refutation. It doesn't even rise to the level of being a good joke. And it is exactly the same, I think, if one's response is instead: "There is a physical reality, but it is radically different from what you always thought; in particular, your knobs on experimental equipment and instrument readings are nowhere to be found there; instead these things only exist in your mind, as, in effect, hallucinations." In both cases, what is proposed is (at best) not in any sense a locally causal account of our empirical observations, but instead a hopelessly philosophical proposal to which Bell's notions of locality and nonlocality are simply inapplicable and irrelevant.

V. QBISM AND QUANTUM SOLIPSISM

Let us then turn to comparing these strange philosophical views (which nobody endorses) to two more serious physical theories (that serious people do seriously endorse and which indeed are put forward, at least in large part, as ways of trying to elude Bell's non-locality). We begin with the more straightforward case of "QBism", which originally stood for Quantum Bayesianism but has apparently now quantum fluctuated its way into standing for some superposition of Quantum Bayesianism, Bohrism, Bettabilitarianism, or (like "KFC", which formerly stood for "Kentucky Fried Chicken" until it was realized that the word "Fried" was a turn-off to health-conscious consumers) nothing at all. [24] This last is perhaps the most appropriate since one of the key ideas of the theory is that quantum states merely summarize subjective beliefs/expectations about future subjective conscious experiences and hence stand for no objectively existing physical structures, i.e., no*thing*.

To elaborate on this a bit, QBism can be understood as the view that takes a personalist Bayesian interpretation of probability (understood quite broadly) and then applies this interpretation specifically to the probabilities that figure centrally in quantum theory. The result, as mentioned, is a thoroughly subjectivist account of the entire quantum calculus, with a particular emphasis on the need to understand quantum theory as a "single-user theory". That is, the probabilities that I use quantum theory to calculate should always and exclusively be understood as *my* credences that *I* assign to various possible future events in *my* subjective conscious experience. To quote one of the theory's main proponents, Chris Fuchs, answering some famous questions posed originally by Bell:

> "*Whose information?*" 'Mine!' *Information about what?* 'The consequences (for *me*) of *my* actions upon the physical system!' It's all 'I-I-me-me mine,' as the Beatles sang." [25]

To understand QBism, it is perhaps helpful to compare it to the Copenhagen interpretation and its infamous "shifty split" between the unspeakable quantum world and the speakable (and, for Bohr, really-existing) world of directly perceivable, macroscopic, "classical" phenomena. [26] QBism is, in effect, the result of moving this shifty split all the way in, so that the speakable "classical" realm includes just the subjective conscious experiences of the single user using the theory. Everything outside is then on the unspeakable quantum side of the split:

> "In QBism, the only phenomenon accessible to Alice that she does not model with

> quantum mechanics is her own direct internal awareness of her own private experience. This (and only this) plays the role of the 'classical objects' of Landau and Lifshitz for Alice (and only for Alice). Her awareness of her past experience forms the basis for the beliefs on which her state assignments rest. And her probability assignments express her expectations for her future experience." [26]

So, with QBism's thoroughly subjectivist understanding of the quantum state, and the idea that everything outside of one's own immediate conscious experience can only be discussed in terms of quantum states, one is left with no means whatever for saying anything definite about any part of the world external to one's own consciousness (if, indeed, there is such an external world at all).[4]

The proponents of QBism vehemently deny that their viewpoint amounts to literal solipsism [27, 28] and I for one am happy to take them at their word. They insist, for example, there there do really exist both outside physical systems (that our observations, measurements, and other interactions with the world observe, measure, and interact *with*) as well as other conscious agents with whom we can communicate (even if, due to the limitations of ordinary language, only imperfectly and approximately).

But at the end of the day, these external systems and other agents play no real role in the theory: to the extent that they can be described at all, they are described only in terms of quantum states. And these, it is insisted, do not provide anything like a direct, realistic description of those outside systems but instead have the exclusively instrumentalist/subjectivist role described above: they summarize our subjective beliefs and expectations about our own future subjective conscious experiences (which may be *as if of* external systems, but need not actually be *of* any such things at all):

> "[W]hen an agent writes down her degrees of belief for the outcomes of a quantum measurement, what she is writing down are her degrees of belief about her potential *personal* experiences arising in consequence of her actions upon the external world." [29]

Let us see how these ideas play out in the QBists' own writings about Bell-type non-locality.

It is worth noting, to begin with, that the QBists' polemics against non-locality are often not directed at Bell's theorem *per se*, but are instead directed at something like the following argument:

> According to ordinary quantum mechanics, the wave function should be taken seriously (as a "beable", to use Bell's terminology) as representing some external physically real thing. But then the collapse postulate immediately implies non-locality: for example, if a single particle is prepared (as in the "Einstein's Boxes" [30] setup) with support in two well-separated spatial regions, looking to see if the particle is present in one of those regions immediately changes the wave function (and hence the physical state) in the other region, no matter how distant. Therefore ordinary quantum mechanics is non-local.

Now that is an important argument, to be sure, and it has some significant overlap with the EPR argument described earlier. But *at most* this argument would establish only that one particular candidate theory – "ordinary quantum mechanics" – is non-local... a radically different conclusion than the one Bell claimed to have established (and, in my judgment did establish), namely, that *any* theory that reproduces QM's empirical predictions will have to be non-local.

In any case, this is the argument against which the QBists' polemics against non-locality are often directed. For example, Fuchs and Schack write that QBism's "thoroughgoing personalist account of *all* probabilities .... frees up the quantum state from any objectivist obligations [and hence] wipes out the mystery of quantum-state-change at a distance..." [29] Fuchs, Mermin, and Schack explain that, according to QBism, "The notorious 'collapse of the wave-function' is nothing but the updating of an agent's state assignment on the basis of her experience." [26] Thus, part of the QBist response to assertions of non-locality is simply the rather obvious point that the specific non-locality – posited by versions of quantum theory in which the wave function is a beable which sometimes collapses – can be avoided if we deny "beable status" to the wave function. That is of course correct as far as it goes. But since the above argument is not at all Bell's argument, it doesn't go very far.

But the QBists do also attack the EPR argument (which they correctly recognize as undergirding Bell's non-locality claim) and in particular the infamous EPR reality criterion:

> "The mistake in the 1935 argument of Einstein, Podolsky, and Rosen lies in their taking probability-1 assignments to indicate objective features of the world, and not just firmly held beliefs. Their argument uses the famous EPR reality criterion: 'If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity)

---

[4] For the record, I should perhaps note that QBism also has a more technical, less philosophical, foot in the realm of "quantum information" and has motivated, and developed in parallel with, a very intriguing project of re-formulating quantum mechanics exclusively in terms of *probabilities* rather than the more standard probability *amplitudes*. These developments, however, while interesting, are orthogonal to the issues we are pursuing here. But readers should be aware that, arguably, there is more to QBism than the philosophical parts I mainly address here.

the value of a physical quantity then there exists an element of physical reality corresponding to this physical quantity.' Without such 'elements of physical reality,' there is no basis for their argument that if quantum mechanics gives a complete description of physical reality, then what is real in one place depends upon the process of measurement carried out somewhere else.

"Bohr maintained that EPR's mistake lay in an 'essential ambiguity' in their phrase 'without in any way disturbing.' But for the QBist their error is simpler. Their mistake was their failure to understand, as many physicists today continue not to understand, that p=1 probability assignments are very firm personal judgments of the assigning agent, and nothing more.

"The unwarranted assumption that probability-1 judgments are necessarily backed up by objective facts-on-the-ground – elements of physical reality – underlies EPR's conclusion that if quantum mechanics is complete then it must be (unacceptably to them) nonlocal. It also underlies Bell's original 1964 derivation of the Bell inequalities." [26]

In terms of my earlier recapitulation of the EPR argument, their point here is that "being certain the distant particle will come out, say, 'V' if its polarization is measured" does not (contrary to what I claimed above) imply that this outcome is somehow encoded in the really-existing independent physical structure of that distant particle. Rather, they suggest, it simply means that one holds a "very firm personal judgment" that, if one later experiences a report of that particle's polarization having been measured, it will be a report of its having been measured to be "V".[5]

Now, it may appear that the QBists are here probing a notoriously soft spot in the EPR-Bell reasoning. As they point out, the criticisms of the EPR criterion of physical reality began immediately with Bohr and have never ceased. It is worth pointing out, then, two things. First, as was mentioned above, the actual EPR paper was written not by Einstein but by Podolsky, and Einstein thought Podolsky's text (and in particular the part involving the reality criterion) obscured the essential argument. And it should be understood that when Einstein himself recapitulated (something like) the EPR argument, his reasoning never included the reality criterion. So, soft or not, the reality criterion is something of a red herring. And then second, the truth is that the criterion is not nearly as soft as is often claimed. [19]

But the more fundamental point here, in understanding the QBists' actual basis for rejecting Bell's nonlocality claim, is that the EPR reality criterion really has nothing to do with it. The relevant difference between their view and Bell's is not in how they interpret "p=1 probability assignments" (and in particular whether they regard them as warranting the attribution of a certain particular structure to the state of the distant particle) but rather in whether they regard any such thing as the distant particle as actually, physically existing in the first place. And here the QBists are thankfully quite clear: the distant measurement outcomes (about which we sometimes make p=1 probability assignments) are on the unspeakable quantum side of the shifty split; our only means of referring to them, in quantum theory, is by assigning an appropriate quantum state. But, according to QBism,

> "each quantum state [has] a home. Indeed, a home localized in space and time – namely, the physical site of the agent who assigns it! By this method, one expels once and for all the fear that quantum mechanics leads to 'spooky action at a distance'.... It does this because it removes the very last trace of confusion over whether quantum states might still be objective, agent-independent, physical properties." [25]

That is, the real reason that (according to the QBists) Bell's EPR-based argument fails to establish non-locality is not because there is some subtle logical invalidity in the EPR reality criterion, but rather because quantum theory, properly interpreted (according to the QBists), assigns all real events to the same one place: "the physical site of the agent". And therefore, since

> "No agent can move faster than light: the space-time trajectory of any agent is necessarily time-like. Her personal experience takes place along that trajectory." [26]

it evidently follows that

> "in QBism quantum correlations are necessarily between time-like separated events, and therefore cannot be associated with faster-than-light influences. This removes any tension between quantum mechanics and relativity." [26]

[5] As Fuchs, Mermin, and Schack explain: "[V]erbal or written reports to Alice by other agents that attempt to represent their private experiences are indeed part of *Alice's external world* and therefore suitable for her applications of quantum mechanics." Thus: "A QBist does not treat Alice's interaction with Bob any differently from, say, her interaction with a Stern-Gerlach apparatus, or with an atom entering that apparatus. This means that *reality differs from one agent to another*. This is not as strange as it may sound. What is real for an agent rests entirely on what that agent experiences, and different agents have different experiences." [26], emphasis added.

Indeed it does.

But, of course, the idea of reconciling quantum theory with relativistic locality in this way is completely absurd. The *point*, presumably, of wanting to maintain consistency with relativity is that one regards the evidence supporting relativity's picture of the world – e.g., as a 3+1-dimensional space-time with a certain structure and populated with certain material objects/fields – as conclusively established. But as I explained in the previous section, there is a huge difference between actually maintaining this picture, and simply avoiding having to assert non-locality. The literal solipsist avoids having to assert non-locality, but certainly does not maintain a relativistic picture of the physical world. Same for the brain-in-a-vat theorist or FAPP solipsist.

And, I of course want to suggest, it is the same as well for the QBist. The QBist picture of the world is one of a "single user" floating (with, for some reason I cannot begin to understand the rationale for, velocity strictly less than the speed of light, whatever that even means in this picture) through a void of immaterial nothingness, or, at least, an unspeakable haze. The very question, for example, of whether the structure of the posited space-time matches that of Einstein and Minkowski, makes no sense. The QBist is free to claim that it does, but what connection would this have to any evidence in or out of QBist quantum theory? Such an assertion, from the QBist, is completely arbitrary and unconstrained, like the anti-Copernican brain-in-a-vat theorist's assertion that the evil scientists live on a stationary planet in the center of some crystalline spheres. Why *that*? It could just as easily be claimed that the real space-time through which the "single user" floats is a 2+1-dimensional Galilean space-time, or anything else. Or nothing at all.[6] There is, at the end of the day, no reason the QBist needs to even posit a physically-existing body for the "single user" – just as the person who thinks he might be a brain-in-a-vat soon realizes that he has no actual reason to believe in the existence of brains, including his own. QBism is in this respect equivalent to the brain-in-a-vat scenario, and hence meets my diagnostic criteria for "FAPP solipsism".

Let me finally reiterate that I am not accusing QBists of being solipsists. The QBists claim that they aren't solipsists and I believe them.[7] I think the actual situation is as follows: it's not that the QBists think there is no external reality. Rather, they do think that there is one, but they also think that quantum mechanics in no sense tells us what it is like (and, at present, neither does anything else). Fuchs and Schack, for example, write that

> "Quantum theory is conditioned by the character of the world, but yet is not a theory *directly* of it. Confusion on this very point, we believe, is what has caused most of the discomfort in quantum foundations..." [29, emphasis added]

But Bell had already responded to this kind of suggestion, decades before Fuchs and Schack made it:

> "You might say, Ψ, the wave function is just not a real thing – it is only a way of talking about something else, and then the fact that it has such funny possibilities of superposition and so on is not so disturbing. But if you say it's not real, I will ask: what *is* real in your theory? What are your kinematics? What are the possibilities that you contemplate and which you talk about when you write down a wave function? .... If your wave function is not real, you must tell me what is." [33]

And that brings us back to square one. If the QBist claim is that quantum theory provides, at best, only some very *indirect* sense of what external physical reality is really like (such that, as noted before, we need not understand the collapse of the wave function as providing immediate proof of non-locality) then we are simply endorsing a vague hidden-variables program kind of view: the quantum mechanical wave function doesn't provide a complete description of external physical reality (and perhaps plays no part whatsoever in such a complete description), but there *is* an external physical reality and someday we will hopefully figure out how to describe it. But the problem is then that any such future candidate realistic account of quantum phenomena (assuming it really accounts for real perceived events involving "knobs on experimental equipment, the currents in coils, and the readings of instruments" in a non-solipsistic way) is subject to Bell's theorem. That is, if it genuinely reproduces the (experimentally well-verified) quantum predictions, it will have to be non-local in Bell's precisely formulated (and, from the point of view of relativity, very troubling) sense. That, simply put, is what Bell's theorem shows. So either QBism is really just a very long-winded (but temporary) *distraction* from Bell's conclusion (in which case it is in no sense a counter-example to Bell's claims), or it is offering a genuinely solipsistic account of a single user having conscious experiences that correspond to

[6] Indeed, Mermin writes that "space-time is an abstraction that I construct to organize [my] experiences." [31]

[7] At least, I believe them most of the time. But there are also other times, for example: "Everything experienced, everything experiencable, has no less an ontological status than anything else. You tell me of your experience, and I will say it is real, even a distinguished part of reality. A child awakens in the middle of the night frightened that there is a monster under her bed, one soon to reach up and steal her arm – that *we-would-call-imaginary* experience has no less a hold on onticity than a Higgs-boson detection event would if it were to occur at the fully operational LHC. They are of equal status from this point of view – they are equal elements in the filling out and making of reality." [25] This really only makes any sense at all if "reality" is understood to mean "subjective conscious experiences" – and if there simply is nothing like what is *usually* meant by "reality".

nothing physically real at all. Either way, in my opinion, it ceases to provide anything like a serious challenge to Bell's non-locality claim.

VI. EVERETT AND QUANTUM SOLIPSISM

I am hardly the first person to suggest that QBism is a rather solipsistic view. [23] But (aside from J.S. Bell! [20]) I don't know of anyone who has suggested that Everett's so-called "Many Worlds" version of quantum theory has a similarly solipsistic character. Indeed, Everettism is generally regarded (along with the de Broglie - Bohm pilot-wave theory and the several extant versions of spontaneous collapse theories) as one of the leading candidate *realistic* accounts of quantum phenomena.

I do not exactly think this is wrong. Certainly it is true that extant *proponents* of Everett's theory are motivated by a healthy sort of realism. And it is certainly true that Everettians posit various really-existing physical structures as part of their theory. Nevertheless, I want to argue that – like the brain-in-a-vat kind of scenario – Everett's theory nevertheless has a FAPP solipsistic character and is thus (superficial appearances to the contrary notwithstanding) very similar to QBism as regards its relationship to Bell's theorem.

Let's start with an overview of the Everettian view. Readers will be familiar with the so-called "measurement problem" of ordinary QM, which can be understood as involving a tension between the unitary Schrödinger evolution of the quantum state, and the various measurement axioms which supposedly govern the interactions between the quantum system in question and a separate classical world. Everett's dual insight was (1) that we can avoid any such tension by simply dropping the idea of a separate classical world and letting the whole universe be governed by the unitary Schrödginer dynamics all the time, and (2) that this isn't as crazy as it initially sounds. It initially sounds crazy, of course, because the unitary Schrödinger evolution produces states involving bizarre superpositions of macroscopically-distinct configurations (such as living and dead cats). It is precisely to avoid such apparently-embarrassing states – which don't seem to have any correlate in our empirical observations of the world – that the measurement axioms are introduced. But Everett suggested that there was in fact no conflict here at all: once it is remembered that we are parts of the universe (not outside, God-like observers of it) we can ask what the world would look like, to us, from the inside, if Everett's postulates were right. And – or, at least, so it is claimed – things would more or less look right, due in no small part to decoherence, which renders the macroscopically distinct "branches" of the universal wave function (effectively) causally independent. And so, crucially, an observer who lives in one such decoherent branch will see and experience only the goings-on in his own branch. And so, it is said, the theory gets the (relevant) appearances right after all.

We are here interested in Everett's theory because Everettians often suggest that the theory constitutes a kind of counter-example to Bell's nonlocality claim. The basis of this suggestion is something like the following idea: the unitary Schrödinger evolution of the wave function was always perfectly local; the non-locality of ordinary QM was always only in the non-unitary collapse dynamics; so by simply getting rid of the collapse postulate we get rid of the non-local part and what remains is purely local. Or the same basic idea can be expressed a little more formally as follows: in an appropriately relativistic version of quantum theory, the law governing the time-evolution of the quantum state (that is, the appropriate relativistic generalization of the Schrödinger equation) is perfectly Lorentz invariant, perfectly compatible with the constraints implied by special relativity. And so (since the theory is that law and nothing more) there is simply no suggestion whatever of a violation of relativistic locality.

While it is true, however, that the fundamental dynamical equations that define Everettian QM are impeccably relativistic, this is not really the same thing as the theory being "local", at least not in Bell's sense.[8] Indeed, as soon as one considers Bell's careful formulation of "locality" – involving probabilities that the theory assigns to various local beables – one realizes that several of the formulation's key notions don't apply, or don't apply *straightforwardly*, to the Everettian theory. For example, the theory doesn't assign probabilities to events, at least not in a familiar kind of way. We might, for example, try to use Bell's formulation to assess the locality of some (more ordinary) candidate theory by examining the probabilities assigned by the theory to some event like "Bob's polarization measurement has the outcome 'H' " when various other events are or aren't conditioned on. But in Everett's theory, the probability that Bob's spin measurement has the outcome 'H' is 100% – as is the probability that Bob's spin measurement has the alternative outcome 'V'! (We assume here a general case in which the amplitudes for both events are nonzero.) The point is, according to Everett, everything (that we normally think of as possible) is in fact deterministically guaranteed to occur, in one of the several downstream branches of the universal wave function. So, for example, the probabilities for (what we normally think of as) mutually exclusive events need not sum to unity; some of our basic ideas about what "probability" even means are thus completely inapplicable, and we must tread very carefully.

Here I do not want to get deeply into this particular issue. How to understand probability from an Everettian perspective is a big and controversial subject. I just want to mention here – in passing – that the subject relates

[8] See, for example, Roderich Tumulka's "rGRWf" theory, which is manifestly relativistic yet non-local. [34]

not only to the question of whether and how Everett's theory can be understood as accounting for the usual probabilistic rules of quantum mechanics (which is the context in which the subject most often arises), but also to the question of whether and how Everett's theory can be described as "local" (in Bell's or some other sense). In particular: to whatever extent Everettians fail to provide a clear and compelling explanation of the meaning of "probability" in the many-worlds context, Bell's notion of locality will remain inapplicable and it will remain invalid to claim that the many worlds theory provides a local (in Bell's sense) explanation of the Bell inequality violations.

A. Wave-Function Realism

But here I instead want to mainly focus on another concept that appears in Bell's formulation of "locality" but which is problematic from the point of view of Everett's theory: local beables. Bell's "locality" is at root a statement about what a theory says about (its) local beables. But in Everett's theory the only beable – the only thing posited as a real physical existent – is the quantum state of the universe. And (however exactly one understands it) this is not a local beable, i.e., in Bell's description, something that attributes physical properties to localized regions (for example, points) in ordinary, three-dimensional, physical space and time. The quantum state can be understood as a (moving) point (or ray) in a very abstract and very high-dimensional Hilbert space. Or perhaps instead as a complex-valued field in some (different, but still abstract and high-dimensional) configuration space. But neither of these representations provides, in any straightforward or clear sense, a description of particles, fields, strings, or any other type of physical "stuff" in (3-dimensional) physical space and time.

Within the school of Everettian thought known as "wave function realism" (according to which it is the universal wave function – the quantum state in position-representation – that provides a kind of direct and complete description of the physical state of the universe), it has been suggested that local beables might be *emergent* rather than fundamental. [36] In such a theory, the fundamental description of the world is in terms of a complex- (or spinor-) valued *field* on a 3N-dimensional space.[9] The fundamental description is simply not of a world of physical "stuff" in 3-dimensional physical space; in Bell's terminology, the theory posits no local beables at all. But, according to the emergence view, one may nevertheless be able to *find* local beables – and in general a familiar-looking 3-dimensional physical world – in the theory, much as one is able to find cats and trees and planets (and even haircuts) in a classical theory whose fundamental ontological posits include only (say) a number of mass points: the cats and trees and planets can be understood as being *made of* the posited mass points. Similarly, in the Everettian case, can we not perhaps understand cats and trees and planets (and hence the kinds of things that are referenced in Bell's formulation of locality) to be *made of* the universal wave function in an analogous manner?

David Wallace and Chris Timpson, for example, write that

> "Three-dimensional features will emerge [from the posited goings-on in the fundamental, high-dimensional space] as a consequence of the dynamics (in large part due to the process of decoherence).... each of the decohering components will correspond to a system of well-localised (in 3-space) wavepackets for macroscopic degrees of freedom which will evolve according to the approximately classical laws displaying the familiar three-dimensional symmetries, for all that they are played out on a higher dimensional space." [35]

And Wallace similarly directly attacks the idea that three-dimensionality must be somehow fundamental rather than merely emergent:

> "How do we know that space is three-dimensional? We look around us. How do we know that we are seeing something fundamental rather than emergent? We don't; all of our observations ... are structural observations, and only the sort of a prioristic knowledge now fundamentally discredited in philosophy could tell us more." [36]

It is my impression, though, that the proponents of Everettism have simply not appreciated the profound difficulty associated with the claim that the three-dimensional world could emerge from (something like) a single field on a much-higher-dimensional space.

To be sure, it is possible for there to be a structural isomorphism between two such realms. For example, the mathematical description of two beads on a straight wire (a system with a two-dimensional configuration space) might be perfectly isomorphic to the mathematical description of a single pool ball bouncing around a square pool table. But just as surely, such mathematical isomorphism does not imply that, any time a pool ball is bouncing around on a pool table, there must exist – in addition – two beads on a wire somewhere. Universes composed of (or including) beads on a wire do not in any sense "emerge" from the dynamical playing-out (in our real universe) of pool-ball-on-table systems. At least, there is not the slightest bit of evidence to suggest that this happens.

[9] or a space of even higher, perhaps infinite, dimension if one considers relativistic field theories rather than N-particle quantum mechanics

Or consider another example. A (three-dimensional!) box contains N gas molecules, each of which (let's say) has at each moment a well-defined position and momentum. The physical state of the collection of N molecules can thus be represented, mathematically, as a single point in a $6N$-dimensional phase space. This works as a representation because there is a perfect isomorphism between the position of a single point in the $6N$-dimensional space, and the positions and velocities of N points in the three-dimensional space. But does this mean that, in any sense, an actual physical particle moving in a new $6N$-dimensional physical space *emerges* from the (let's assume) really-existing collection of N molecules? Of course not. To think so is simply to reify what is in fact merely an abstract representation. And it would be exactly the same if it were (strangely) the one-particle-in-$6N$-dimensional-space story that were regarded as fundamental, and the three-dimensional-box-of-gas that were supposed to "emerge".

About just this kind of case, Wallace and Timpson seem to take a different view:

> "if the N-particle story were empirically adequate (which it isn't, of course) then so would the one-particle story be. For that story (by construction) is isomorphic to the N-particle story, and the emergence or otherwise of higher-level ontology from lower-level theories depends, to our minds, primarily on the structure of those theories and not on their underlying true nature (whatever that is). On the one-particle theory, three-dimensional space would turn out to be emergent, but it would be no less real for that..." [35]

But I simply do not understand how to make sense of this view, if the "emergence" of a three-dimensional space containing three-dimensional objects is supposed to mean that that space and those objects really come into existence. It is easy to understand how high-level macroscopic structures such as cats and haircuts (which involve specific patterns of three-dimensional stuff) can emerge from some fundamental ontology of three-dimensional stuff (particles, fields, whatever). But a cat, or a haircut, is an essentially three-dimensional pattern, and I don't understand how a physically real three-dimensional pattern of any kind can emerge from an underlying reality in which the physical stuff lives in a very different (much higher-dimensional) space. We could have a higher-dimensional pattern that is somehow isomorphic to the three-dimensional pattern, but this is not the same as having a really-existing three-dimensional pattern.

But perhaps my incomprehension regarding such "trans-dimensional emergence" is misplaced and irrelevant. Perhaps, that is, what is being suggested (when it is suggested that three-dimensional things like cats and planets and pointers and haircuts can emerge, for example, from a wave-function-realist ontology) is only that the *appearance* of cats and planets and pointers and haircuts will emerge. That is, perhaps the claim is merely that, for conscious inhabitants of the posited kind of world, it will *seem* as if they live in a three-dimensional world inhabited also by cats, planets, pointers, and haircuts.

Let me immediately concede that this view – what we might call "appearance-emergence" – is indeed rather plausible. The idea is that, although nobody would claim to understand the process in detail, and although many deep and profound ontological mysteries remain, we are all more or less comfortable with the idea that conscious experience somehow *emerges from* the complicated physical processes occurring in a living, functioning (three-dimensional) brain. And further, we are all reasonably confident that there is in this regard nothing unique about the specific "wet physical stuff" that happens to comprise human brains. So if, for example, we could replicate the requisite *brain structure* in some other physical embodiment (say, an enormously-complicated network of silicon-based transistors) consciousness would again emerge. And in particular, it seems quite likely that if the two networks (of wet neurons on the one hand, and silicon transistors on the other) were perfectly structurally isomorphic, then not only would consciousness emerge in both cases, but, it seems quite plausible to suggest, the emergent consciousnesses would in some sense be exactly the same. But from there it is no great leap to the suggestion that in a universe that is radically different from ours – but nevertheless *structurally isomorphic to ours* – not only would the "beings" there be conscious in the same way that we (imagining ourselves to live in a 3-dimensional universe) take ourselves to be, but indeed the conscious experiences of those beings would be identical to ours.

It is in this sense, I think, that the Everettians claim – plausibly in my opinion – that an Everettian universe (consisting of, say, a field in a very high-dimensional space evolving according to Schrödinger's Equation) would contain conscious beings whose conscious experiences perfectly matched ours – including, evidently, the experience of inhabiting a three-dimensional world populated by cats, trees, planets, etc. Thus, while I don't think the Everettian can plausibly claim that local beables in fact emerge (trans-dimensionally) from a wave-function-realist picture, I think he can plausibly claim that the subjective conscious experience of local beables might so emerge. In effect, the inhabitants of an Everettian universe would be deluded into thinking that they lived in a 3D world just like the one we take ourselves to inhabit. And so, according to the Everettians, for all we know, we might very well be those inhabitants of an Everettian universe.

And that, as far as it goes, is probably correct. There is no specific element of anyone's subjective conscious experience – there is no specific evidence – that one could point to that would show that we are *really* made of particles and fields moving and interacting in a 3-dimensional space, as opposed to being *really* made of field-stuff in a much higher-dimensional space. But then, there is not supposed to be any specific element of subjective conscious experience that one could point to, either, to prove that we are not *really* brains-in-vats. And the parallel there should by now be entirely clear. The "appearance-emergence" Everettian is offering a worldview in which the *true* physical world is radically different from the one we ordinarily take ourselves to be perceiving. In this view, every aspect of our ordinary perceptual experience (and hence also everything that is based on it) is fundamentally delusory, just as in the FAPP solipsist brain-in-a-vat scenario. And so I think all of the reasons we have for not taking those FAPP solipsist ideas too seriously (even though, as Bell noted, they cannot exactly be refuted) apply as well to Everettianism.

The Everettian will of course want to insist that his worldview – his wave-function ontology – is totally unlike the brain-in-a-vat worldview since wave-functions (unlike the hypothesis that one is a brain-in-a-vat being controlled by evil scientists) enjoy considerable empirical scientific support. But I think a little reflection will reveal that the positions are not actually so different. The Everettian here has precisely the same type of empirical grounds for believing in wave-functions, as the brain-in-a-vat theorist has for believing in brains – namely, inference from a wealth of perceptual experiences which, according to the belief in question, were fundamentally misleading.

My point is not simply to tar Everettism by association with solipsism. That is, my point is not to say: "Everettian QM has a FAPP solipsistic character; this makes it *crazy*, so we shouldn't take it seriously." I actually do take it quite seriously. My point is entirely different: not that Everettian QM is crazy, but that the claim that it is a *local* theory – some kind of counter-example to Bell's claim that non-locality is required to explain certain observed facts – is at best empty and misleading and obscurantist rhetoric. The truth is that Everettian QM (at least in the wave-function realist version, interpreted via "appearance-emergence") is neither local nor nonlocal in Bell's sense. What it posits as the fundamental nature of physical reality is so completely different from the kind of picture Bell assumes, that his definition of "locality" fails entirely to apply in any meaningful way at all.

B. Space-time State Realism

Recently, though, another way to understand the ontology of Everettian QM has been proposed, and it is worth examining that as well, lest our diagnosis be too broad or too hasty. I have in mind here Wallace and Timpson's "space-time state realism" (SSR) which, I think it is fair to say, they propose precisely to avoid the type of difficulty one gets into when one's theory contains no local beables:

> "Our claim, in essence, is that thinking about quantum mechanics in terms of a wavefunction on configuration space is rather like thinking about classical mechanics in terms of a point on phase space. In both cases, there is a far more perspicuous way to understand the theory, one which is connected to spacetime in a more direct way." [35]

And in particular SSR is supposed to help clarify the status of Everettian QM vis-a-vis nonlocality:

> "Our project ... might seem simple. Does Everettian quantum mechanics violate Local Action? And does it violate Separability? But things are not quite that easy: even saying what *is* the state of a given physical region in quantum theory requires us to have a more solid grasp of the physical reality that 'the quantum state' represents than is available from its abstract, Hilbert-space definition. [We need to first] get a firmer grip on just what the physical world is like according to Everettian quantum physics." [37, pp. 294-5]

Wallace, that is, seems to agree with Bell that one can only make meaningful claims about locality or nonlocality after clearly specifying a theory's ontology of local beables.

The main idea of SSR is to take the ontology of the theory to be explicitly three-dimensional: a set of properties of space-time regions, captured in particular by the set of *density operators* over all space-time regions. Wallace and Timpson suggest that this ontology is comparable to that of classical electromagnetism, whose fields can also be understood as describing properties of spacetime regions (here, points). Of course, the idea of *operators* as local beables is a little unusual. Operators don't take *values* the way that, say, an electric field does... but as Wallace argues, maybe objecting to operators-as-beables is simply based on the novelty and unfamiliarity of the idea, rather than any actual problem with it. Maybe. In any case, whereas it is completely clear how cats and planets and trees can be made of the local beables in (for example) Bohmian mechanics (namely: they are simply cat-, planet-, and tree-shaped constellations of particles), it remains at best obscure how to understand spacetime state realism's local density operators as providing an image of the familiar everyday three-dimensional world.

But let us focus on a different point, more relevant to our overall project of questioning the Everettian claim to provide a *local* explanation of the EPR-Bell correlations. To be clear, the advocates of spacetime state realism concede that the theory is nonlocal in a certain sense: it involves *non-separable* state descriptions. This non-separability plays a crucial role in the theory's (purported) ability to account for the EPR-Bell correlations. But, the advocates stress, non-separability is very different from the *dynamical nonlocality* whose necessity Bell claimed to have established. And in particular the space-time state realists claim that their non-separable but (dynamically) local theory is fully compatible with relativity in a way that a dynamically nonlocal theory never could be.

Let us examine these claims in some detail. First, let us understand exactly the nature of the theory's non-separability. The reduced density operators that spacetime state realism posits as (something like?) local beables, are formed by taking the density operator of the universe and doing a partial trace over the degrees of freedom from outside the spacetime region in question. The density operator for a sub-region (of a given, larger region) can thus always be formed from the density operator of the larger region. But we cannot re-compose the density operator for a larger region, even from the density operators for a set of sub-regions that jointly cover the entire larger region.

More concretely, in the standard kind of EPR-Bell scenario in which spatially-separated observers Alice and Bob make measurements (along the same axis, say) of the polarizations of each photon from an appropriately polarization-entangled pair, the density operator for the large region (comprising both Alice's and Bob's labs) will be a mixture of operators corresponding respectively to "Alice and Bob both measure the polarization to be 'H"' and "Alice and Bob both measure the polarization to be 'V'." From this, we may compute the density operator describing the state of Alice's lab alone: it is a mixture of (operators somehow supposedly representing) "Alice measures the polarization to be 'H' " and "Alice measures the polarization to be 'V'." And similarly for the density operator describing Bob's lab alone. The point is then that, from the descriptions of Alice's and Bob's labs separately, we could not recover the fact that their outcomes are perfectly correlated. (In the Everettian picture, this perfect correlatedness of course means that the universe has split into two branches, one in which both Alice and Bob measure the polarization to be "H", and one in which they both measure the polarization to be "V"; there is no branch in which Alice's and Bob's outcomes are different.) This is the sense in which the theory involves non-separability: there are certain facts pertaining to larger spacetime regions (like the perfect correlatedness of Alice's and Bob's polarization measurements) which cannot be decomposed into sets of facts pertaining to the sub-regions that jointly compose the larger region.

I think it is worth pointing out here that all of this seems to me to render these would-be local beables rather pointless. Basically this non-separability means that any description of the state of several limited spacetime regions will be decidedly *incomplete* unless we also include a description of the state of the larger region that comprises them. But once we do that, the states of the sub-regions are entirely redundant. Why bother positing them, as separately-real existents, at all?[10] And of course this argument can be immediately repeated: if it is pointless to posit separate "spacetime states" for Alice's and Bob's labs (but instead only the larger region comprising both labs), then it is equally pointless to posit a "spacetime state" for that region comprising both labs; we should instead just posit the spacetime state of an even larger region comprising both labs and, say, some cat somewhere; and so on, until we realize that the only "spacetime state" we really need to posit at all is the one for the ultimate spacetime region – the universe as a whole. But the density operator for the universe as a whole is mathematically equivalent to the quantum state of the universe as a whole. Which leaves us right back where we started: Everettianism as positing the quantum state (perhaps in something like "position representation", i.e., wave function realism) and running into serious trouble because of the lack of local beables.[11]

It could perhaps be argued that this criticism of spacetime state realism is unfair. A similar argument, for example, could be leveled against any theory in which the local beables are *functions* of the quantum state. For example, in the theory "GRWm" (which posits a universal wave function obeying the non-linear GRW evolution equation [38]) the role of local beable (out of which cats and planets and trees are supposed to be made) is played by a "mass field" $m(\vec{x},t)$ which is simply defined in terms of the universal wave function $\Psi(\vec{x}_1,\vec{x}_2,...,\vec{x}_N,t)$. If Ψ is given at some time t, then the mass field m can be computed. So, one might argue, what is the *point* of positing the mass field as a separate existent? Isn't it redundant? In my opinion, the answer here is: no. The mass field is not redundant, because of the crucial role it plays (in the context of this particular theory) in providing an image of the familiar three-dimensional world of everyday perception, including not only the cats and planets and trees I keep mentioning, but also the instrument pointers whose

---

[10] Interestingly, the proponents of SSR sometimes seem conflicted about whether the local density operators constitute additional, separately-posited ontology (beyond the quantum state), or should instead be regarded as merely parts of, or perspectives on, the quantum state (which would constitute the entire ontology by itself). For example, Wallace writes: "Everettian quantum mechanics reads the quantum state literally, as itself standing *directly* for a part of the ontology of the theory. To every different quantum state corresponds a different concrete way the world is, and the quantum state *completely* specifies the ontology." [37, p. 295] So... which is it? Is the quantum state only a *part of* the ontology (with the local density operators presumably filling the ontology out)? Or does the quantum state (by itself) *completely* specify the ontology? It is not really clear. My point here is just that the ontic redundancy noted in the main text may explain the proponents' ambivalence.

[11] Wallace expresses the trouble this way: "Note that if ... we were to treat the Universe just as one big system ... then we would only have a single property bearer (the Universe as a whole) instantiating a single property (represented by the Universal density operator), and we would lack sufficient articulation to make clear physical meaning of what was presented." [37, p. 299-300]

positions ultimately constitute so much empirical data. So after all it does I think make an important difference that spacetime state realism's would-be local beables are (qua operators on an abstract mathematical state space) so difficult to understand as describing, in any straightforwardly comprehensible way, a (set of) familiar-looking three-dimensional world(s).

Despite the apparent impossibility (or perhaps, more generously, obscurity) of understanding spacetime state realism's density operators as providing an image of the familiar physical world of ordinary perception, it should be admitted that – just as with wave function realism – there is a kind of structural isomorphism between (one "part" or "branch" of) the reduced density operators and the ordinary 3D physical world. So from a certain "functionalist" point of view, both versions of Everettism may perhaps be considered viable, at least in the FAPP solipsist sense I discussed above for the case of wave-function realism. But instead of repeating that kind of analysis, here I want to push a bit farther into the claim that SSR's (would-be) local beables allow a clear diagnosis of dynamical locality. So let us temporarily set aside the worries about the adequacy of the SSR local beables and grant to the Everettian spacetime state realist more or less everything he wants to claim. I will then argue that it is *still* highly questionable whether the theory provides a genuinely (dynamically) local account of the empirically observed correlations.

Let us grant, in particular, that SSR's reduced density operators provide an unproblematic slate of local beables in terms of which we can find, in the world posited by the theory, events such as "Alice measuring her photon's polarization to be 'V'." Let us also grant that we can apply Bell's notion of "local causality" (i.e., dynamical locality) to the Everettian theory by replacing the probabilities (that appear in Bell's formulation) by the "branch weights" that play a somewhat analogous role in many worlds type theories. In particular, let us amend Bell's formulation as follows (see again the earlier figure):

> "A [many-worlds-type] theory will be said to be locally causal if the [branch weights] attached to local beables in a space-time region 1 are unaltered by specification of values of local beables in a space-like separated region 2, when what happens in the backward light cone of 1 is already sufficiently specified, for example by a full specification of local beables in a space-time region 3...." [18]

Does the Everettian SSR theory come out as "locally causal"? This very much depends on exactly what set of facts (posited by the theory) one regards as included in the local beables. For example, let us take the event in question (in region 1) to be "Alice's polarization measurement has outcome 'H'." The branch weight of this event is (let's say, assuming an obvious kind of setup) 0.5. And that will be true independent of whether one also considers the state of Bob's laboratory. Consider in particular the situation where Bob decides, at the last minute, whether to measure his photon along the same axis as Alice, or an orthogonal axis. In the former case, the density operator for the larger region (jointly describing both Alice's and Bob's labs) will be a 50/50 mixture of "HH" and "VV" (i.e., "Alice's outcome is 'H' and Bob's outcome is 'H"' and "Alice's outcome is 'V' and Bob's outcome is 'V"'). And so the total weight of branches in which "Alice's outcome is 'H"' will be 0.5. Whereas in the latter case, where Bob chooses instead to measure along an orthogonal axis, the density operator for the larger region will instead be a 50/50 mixture of "HV" and "VH". And so, again, the total weight of branches in which "Alice's outcome is 'H' " will be 0.5. In short, the branch weight of this (localized) event is unaffected by including the information (from Bell's region 2) about Bob's choice of measurement axes. And so one would apparently conclude that the theory is locally causal, in the modified Bell sense.

But consider another fact pertaining to Alice's measurement to which we can apply the modified Bell locality criterion: whether the (descendant of) Alice who ends up seeing the "H" outcome is in the same universe (or "branch") as the Bob who ends up seeing the "H" outcome. If Bob happens to choose to measure along the same axis as Alice, the descendant of Alice who observes the "H" outcome will definitely end up in the same universe as the descendant of Bob who observes "H". Whereas if Bob instead happens to choose to measure along the orthogonal axis, the descendant of Alice who observes "H" will definitely *not* end up in the same universe as the Bob who observes "H". The weight of branches in which the "H"-observing Alice is in the same universe as the "H"-observing Bob thus very much depends on Bob's choice of measurement axis: it is 1.0 for one of Bob's possible (free) choices, and zero for the other.

Does this constitute a clear violation of the modified Bell locality criterion and hence a demonstration that the SSR version of Everettian QM is (dynamically) non-local after all? I'm honestly not sure. It depends on whether one allows something like "The Alice-descendant who observes 'H' is in the same universe as the Bob-descendant who observes 'H' " as a statement *about* Alice, i.e., as a local beable pertaining to Bell's spacetime region 1. And I can see arguments cutting both ways there. On the one hand, this is somehow clearly a statement about one of Alice's descendants who is present in region 1. On the other hand, you couldn't see whether this was true or not if you *only* looked at the reduced density operator pertaining to region 1. The fact in question is thus somehow clearly a non-local relational fact pertaining to both regions 1 and 2. And when one recognizes that, it becomes less surprising that Bob's free choice (which occurs in region 2) could influence it.

However, I do not think it would be right to simply dismiss the example as failing, after all, to show that there is some non-locality in the theory. Of course it is true that no non-locality is implied by the fact that Bob's free choice in region 2 can affect the joint state of regions 1 and 2. But remember: not only is the fact in question here (namely, whether the Alice-descendant who observes "H" is in the same universe as the Bob-descendant who observes "H") not present in region 1 alone – it is also not present in region 2 alone. So it is wrong to think that what Bob's free choice is causally influencing is merely the state of region 2. It is instead causally influencing the joint state of regions 1 and 2. And if a free choice in one region can causally influence the physical state of another region, not all of which is confined to the future light cone of the first region, that does start to sound suspiciously like a violation of relativistic causal structure – even if (because of the "non-separability") it is impossible to pinpoint a particular space-time location for the effect.

Coming at this same point from another direction, Wallace wants to suggest the following kind of picture: Alice's measurement triggers a branching event which propagates outward from her lab at (roughly) the speed of light; similarly, but independently, Bob's space-like separated measurement triggers another branching event which propagates outward from his lab at the speed of light; it is only when/where these two outwardly-propagating branchings come to overlap (that is, it is only in the intersection of the future light cones of the two measurement events) that the different branches get "connected up" in such a way as to address, for example, whether the Alice-descendant who observes "H" is in the same universe as the Bob-descendant who observes "H". [37, p. 307] This has the superficial appearance of providing a perfectly locally causal mechanism for the kinds of correlations that are in question here (e.g., whenever Alice and Bob measure along the same axis, they find, when they compare notes later, that their outcomes are always identical).

The grain of truth in this story is that, indeed, these relational or correlational facts only become manifest in the strictly local beables in the overlapping future light cones of the individual measurement events. But – and this is the crucial point here – those facts are nevertheless quite real well before this time, as is clear from simply considering the state of the region composed jointly of 1 and 2.

Wallace of course recognizes that, in this EPR kind of case, the world-splitting events triggered by Alice's and Bob's measurements are not independent (i.e., he recognizes that there already exist non-locally-instantiated facts about which of Alice's and Bob's descendants are in the same worlds as each other):

> "Nor is this to be expected: ... in Everettian quantum mechanics interactions are local but states are nonlocal. The entanglement between the particle at A and the particle at B is a nonlocal property of [the joint region] $A \cup B$. That property propagates outwards, becoming a nonlocal property of the forward light cone of A and that of B. Only in their intersection can it have locally determinable effects – and it does, giving rise to the branch weights which, in turn, give rise to the sorts of statistical results recorded in Aspect's experiments and their successors: statistical results which violate Bell's inequality." [37, pg. 310]

I would summarize the situation differently: I think the SSR theory only looks dynamically local to the extent that one does not take its many worlds character sufficiently seriously. In particular, SSR defines the local ontology as the sum total of what's happening in a given spacetime region across all the worlds, rather than taking the distinct worlds (and in particular their distinctness from one another) seriously.[12] If one did take the distinct worlds more seriously – by giving them pride of place in the local ontology – it would become immediately clear that the theory is as non-local as it can make sense for a many-worlds theory to be: which universe Alice's descendants are born into (namely one in which Bob's measurement came out "H" or instead one in which it came out "V") depends, non-locally, on Bob's space-like separated choice of which axis to measure his photon's polarization along.[13]

This point is closely related to the point I raised above about the local beables (in SSR) being ultimately redundant. Much more is *physically real*, according to SSR Everettianism, than is manifest in the reduced density operators of localized spacetime regions. Wallace's claim is essentially that if one assumes these density opera-

---

[12] There is an interesting parallel here to non-Everettians who similarly define the local state as a reduced density operator and argue that there is no non-locality since distant interventions do not affect the local state. [4] In a non-many-worlds context, the density operator is really just a catalog of the expected statistics for an ensemble of identically-prepared systems, and its not changing (as a result of distant interventions) is equivalent to the familiar inability to send superluminal signals. But it was never a violation of such "signal locality" that Bell claimed to have established. The point here is just that Wallace's defense of the supposed dynamically local status of SSR Everettianism seems to commit a kind of many-worlds-analog of this standard fallacy of switching the meaning of "locality" from Bell's "local causality" to "signal locality".

[13] Note the subtly solipsistic turn in Wallace's discussion of this: "From the perspective of a given experimenter, of course, her experiment *does* have a unique, definite outcome, even in the Everett interpretation. But Bell's theorem requires more: it requires that from her perspective, her distant colleague's experiment also has a definite outcome. This is not the case in Everettian quantum mechanics – not, at any rate, until that distant experiment enters her past light cone. And from the third-person perspective from which Bell's theorem is normally discussed, no experiment has any unique definite outcome at all." [37, pg. 310] The point is that the locality or nonlocality of the theory depends on whether the theory says that what happens to Alice depends on what Bob did, or does not so depend. It has nothing to do with what things look like "from her perspective". Wallace here conflates the question of whether, according to the theory, Bob's actions influence Alice, with the totally irrelevant question of when and where Alice might *find out* about Bob's actions.

tors (for the localized regions) capture everything that is actually going on in space-time, then there is no nonlocality. That may be true. But if one instead remembers the facts that are, according to the theory, physically real, but which are not captured by the local density operators (for example, the way that the different terms/branches at different locations *correlate* with one another), the suggestion that the theory is local looks very suspicious.

All of that said, though, let me be the first to confess that all of this is less than perfectly clear and compelling. Recall Bell's under-appreciated point:

> "If local causality in some theory is to be examined, then one must decide which of the many mathematical entities that appear are supposed to be real, and really here rather than there." [2]

Wallace and Timpson's SSR is a step in the right direction for Everettianism in the sense that it represents a tacit acknowledgement that local beables are required in a theory that purports to give a realistic account of empirical (i.e., observed) reality. But the truth is that the SSR ontology remains far too obscure to allow any clear and unambiguous analysis of the sort I'm attempting above.

To review, there are several issues. First, the would-be local beables are abstract mathematical operators that act on vectors in a Hilbert space. I appreciate and accept Wallace's point that

> "[t]here need be no reason to blanch at an ontology merely because the basic properties are represented by such objects [namely, abstract operators]: we know of no rule of segregation which states that, for example, only those mathematical items to which one is introduced sufficiently early on in the schoolroom get to count as possible representatives of physical quantities!" [37, p. 299]

But polemics against rejecting the proposed ontology don't really help. What's needed is a clear account of how the proposed ontology can be understood: what sort of three-dimensional thing do these density operators represent, and how exactly does the representation work? Without answers to such questions, it is simply not clear what the theory is supposed to be about, i.e., what, according to the theory, is supposed to be real.[14]

The second issue I mentioned above is that the SSR ontology of local beables seems somewhat like a metaphysical afterthought. Not only are the local beables mathematically redundant (in the sense that the density operator of any large region determines the density operator of all its sub-regions, but not vice versa), but it is certainly not possible to reformulate the theory mathematically in terms of the local density operators. In short, it simply does not appear that Everettian SSR is a theory *about* local density operators. Instead, the theory looks more like (because, in fact, it is) a theory that is fundamentally *about* the universal quantum state, with some colorful but questionable and probably superfluous window dressing of an only vaguely local beablish character.

At the end of the day, then, I think the only currently-available way to make any sense of Everettian quantum theory is that it describes a *physical* world radically different from the familiar three-dimensional world of everyday perception – for example, a world consisting of a dynamical field evolving in a very high-dimensional space (i.e., the wave-function realist picture). Nobody has proposed a comprehensible way of understanding how the familiar three-dimensional world of ordinary perception might be represented by, or might somehow emerge from, the universal wave function. What is (at least somewhat) plausible, though, is that the structural isomorphisms between certain dynamical degrees of freedom in the evolving wave function, and human neurophysiology, might imply that "information processing agents" in an Everettian world should have conscious experiences identical to those usually assumed to emerge from human neurophysiological processes. That is, the beings in an Everettian world might be expected to have conscious experiences just like ours, i.e., these beings might be deluded into thinking that they are humans who live in a three-dimensional world populated as well by cats, trees, and planets.

Extant Everettianism is, in that sense, an elaborate brain-in-a-vat proposal, and hence a clear case of what I have called "FAPP Solipsism". It is not, I think, the case at all that it provides a legitimate counter-example to Bell's claim that non-local dynamics is required to explain the (apparently) observed predictions of quantum mechanics. In fact, as has been pointed out before, it denies that those predictions, as usually understood, even actually occur. [39] But the radical extent of this denial has not been sufficiently appreciated. In fact, according to Everett's theory, nothing like the three-dimensional world containing Alain Aspect, his laboratory, and the pointer positions (with their famous Bell-inequality violating correlations) actually exists. Instead we have, at best, a kind of delusional appearance of such things, in the minds of beings in a radically different kind of physical universe – a story that, even on its own premises, it is very difficult to take seriously.

[14] And, for the record, "what's real is something that can be accurately and completely represented by density operators" does not suffice here. One might as well just forget about the SSR ontology and claim that "what's real is something that can be accurately and completely represented by the quantum state". The very same thing that motivates Wallace and Timpson to propose SSR in the first place – namely the inadequacy of the kind of answer quoted in the previous sentence – should make it clear that SSR remains, at best, inadequate.

VII. CONCLUSIONS

Critics of Bell's claim to have established the reality of non-local causal influence in nature often just erroneously assume that his analysis *begins* with two premises: locality (i.e., no faster-than-light causal influences) and deterministic "hidden variables". Such critics are simply ignorant of Bell's actual reasoning, which begins not with an assumption of determinism but rather with a proof (due, roughly, to EPR) that, already in the case of the narrow subset of correlations considered by EPR, locality *requires* determinism. Bell's careful formulation of locality, reviewed in Section III above, provides a helpful corrective here, in so far as it allows the argument "*from locality to* deterministic hidden variables" [13] to be made in a more formal and rigorous way. (See, e.g., Ref. [9] for details.)

But Bell's formulation of locality is also quite helpful in confronting another rather different category of critics – namely, those (like the QBists and Everettians) purporting to offer fully-worked-out versions of quantum theory which elude the supposedly-necessary kind of nonlocality. What is helpful, in particular, is Bell's insistence that the very idea of locality can only be formulated in terms of "local beables" – i.e., his insistence that a theory must include, in its ontology, appropriate building blocks for directly observable macroscopic things such as "the settings of switches and knobs on experimental equipment ... and the readings of instruments." [15] What is the sense of this alleged "must"? The point is that directly-observed (macroscopic) physical objects, including the pointers which register the outcomes of physics experiments, consist of such ontological facts. A theory in which such facts are not *physically real* – but are instead some kind of hallucination or delusion – necessarily gives only a solipsistic account of observed phenomena. This means that it doesn't actually account for the physically-realized outcomes that have been observed in various experiments, but instead only purports to account for the (delusional) subjective appearance, in consciousness, as if of such outcomes.

We discussed two concrete examples of such FAPP solipsistic accounts.

The essence of the QBist account is an explicit and systematic denial that the theory posits any ontology at all. This denial is somewhat undercut by the proponents' protests against charges of solipsism, as well as their somewhat absurd suggestions that (what they confusingly call) "reality" (for, recall, the single user of the theory) propagates along a well-defined, sub-luminal trajectory through physical space. But at the end of the day the theory's status is perfectly clear: everything it purports to be about, and everything we could ever hope to possibly assert on its basis, occurs inside the consciousness of some lonely agent. With no physical ontology, the solipsistic character is clear – and with it the complete and total inapplicability of concepts like "local" and "nonlocal". So while it is certainly the case that QBism provides a way of trying to understand quantum mechanics that avoids any commitment to spooky, anti-relativistic action-at-a-distance, it is totally uninteresting as a supposed counter-example to Bell's claim that non-local causal influences are required to explain what is observed in experiments. For QBism, qua FAPP solipsism, denies that anything was in fact observed in those experiments and, indeed, denies that any such experiments every actually, physically, took place at all.

Everettian quantum theory, on the other hand, is a little harder to pin down. In its so-called "space-time state realist" formulation, there is at least a nod in the direction of trying to provide an ontology of local beables. But this remains, at best, a work in progress; at present I simply cannot understand how the abstract mathematical operators (posited as "local beables") can be understood as describing the state of three-dimensional matter in the necessary table-, cat-, tree-, and switches-and-knobs-on-experimental-equipment like configurations. (And even if the density operators could be understood as providing an appropriate image of three-dimensional physical reality, it remains far from clear that the theory is actually local in Bell's sense.) The "wave function realist" formulation, on the other hand, provides a perfectly comprehensible ontology: a dynamical field on a very high dimensional ("configuration") space. But in this picture we are, at best, and absent some future explanation of trans-dimensional physical emergence, apparently like the proverbial brains-in-vats whose every subjective conscious experience is a hallucination or delusion. Thus, in the only way I can at present make any sense of Everettian quantum theory, it too has a FAPP solipsistic character and hence cannot really be taken seriously as providing any meaningful sort of counter-example to Bell's arguments.

In the case of QBism, I have no hope whatever that the theory might someday be developed to provide a coherent physical ontology. The denial of any such ontology is just too central to the whole motivation. In the case of Everettianism, on the other hand, I am not so sure. Indeed, there is perhaps already a hint of a more viable strategy for a coherent Everettian ontology in Ref. [40]. So I do not mean at all to claim that I have somehow once-and-for-all diagnosed Everettian quantum theory as solipsistic. My claim is much more reserved: the ontology – the local beables – of Everettian quantum theory remain, to me, so obscure that I cannot yet understand the theory as providing a realistic description of directly observable macroscopic reality. And – *therefore* – any suggestion that the theory somehow provides a counter-example to Bell's nonlocality claim is radically premature.

The lesson of all this, as we celebrate the 50th anniversary of Bell's great achievement, is simply that we cannot forget Bell's insistence on the preconditions of discussing locality, and in particular his point that this concept can only be really *understood* "in terms of local beables". [15] In Bell's view, the local beables of any serious theory have to "include the settings of switches and knobs on experimental equipment ... and the readings of instruments" [15] in the sense of including some physically real stuff out of which such directly observable macroscopic objects can be coherently understood to be *made*. This, at the end of the day, is the only notion of "realism" that Bell assumes. But who could deny this? Surely even the most strident instrumentalist believes in the physical reality of directly observable macroscopic instruments! It seems that the only people who could possibly deny this very elementary sort of "realism" are solipsists and FAPP solipsists. I have suggested that there are several influential camps of FAPP solipsists involved in these quantum foundational discussions. But, revealed as such, these need not (and can not) be taken too seriously.

What Bell established 50 years ago is really quite remarkable: as long as you understand the ordinary world of directly observable macroscopic objects in a standard, non-solipsistic way, there is no way of filling out the microscopic details of the ontology and dynamics such that experimentally observed correlations are accounted for in a purely local way. Contrary to the assertions of many commentators, this leaves us really only two options: we can accept that nonlocality is a real feature of our world, or we can adopt a bizarre, solipsistic type of view that is far, far less palatable and (at best) "local" only in some totally arbitrary and empty sense.

Perhaps in another 50 years, when we celebrate the 100-year anniversary of Bell's discovery, the ridiculous lengths to which one must go to avoid non-locality will finally be widely appreciated?

---

[1] J.S. Bell, "On the Einstein-Podolsky-Rosen paradox," *Physics* **1** (1964) 195-200; reprinted in J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, 2nd ed., Cambridge, 2004.

[2] J.S. Bell, Preface to the First Edition of *Speakable and Unspeakable in Quantum Mechanics*, 1987, *op cit.*

[3] A. Aspect, J. Dalibard, and G. Roger, "Experimental test of Bell's inequalities using time-varying analyzers," *Phys. Rev. Lett.* **49** p. 1804-1807 (1982); G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, "Violation of Bell's inequality under strict Einstein locality conditions," *Phys. Rev. Lett.* **89** pp. 5039-5043 (1998)

[4] Reinhard Werner, "Comment on 'What Bell did'," J. Phys. A: Math. Theor. **47** (2014) 424011 (8pp). See also "What Maudlin replied to," arxiv:1411.2120 and Werner's blog post (and the comments thereon) at http://tjoresearchnotes.wordpress.com/2013/05/13/guest-post-on-bohmian-mechanics-by-reinhard-f-werner

[5] Marek Zukowski and Caslav Brukner, "Quantum non-locality – it ain't necessarily so..." J. Phys. A: Math. Theor. **47** (2014) 424009 (10pp).

[6] A. Aspect, Introduction to the 2nd Edition of *Speakable and Unspeakable in Quantum Mechanics*, by J.S. Bell, Cambridge, 2004

[7] J.S. Townsend, *A Modern Approach to Quantum Mechanics*, 2nd ed., University Science Books, Mill Valley, CA, 2012

[8] S. Goldstein, T. Norsen, D. Tausk, and N. Zanghí, "Bell's Theorem", `http://www.scholarpedia.org/article/Bell's_theorem`

[9] T. Norsen, "J.S. Bell's concept of local causality" *American Journal of Physics* **79** (12) December 2011 pp. 1261-1275

[10] A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.* **47** 777 (1935)

[11] A. Fine, *The Shaky Game*, University of Chicago Press, 1996

[12] D. Howard, " 'Nicht sein kann, was nich sein darf, or the prehistory of EPR: Einstein's early worries about the quantum mechanics of composite systems" in A.I. Miller, ed., *Sixty-Two Years of uncertainty: Historical, philosophical, and physical inquiries intot he foundations of quantum mechanics*, pp. 61-111, Plenum, New York, 1990

[13] J.S. Bell, "Bertlmann's socks and the nature of reality," *Journal de Physique*, Colloque C2, suppl. au numero 3, Tome 42 (1981) pp C2 41-61; reprinted in *Speakable and Unspeakable*, *op cit.*

[14] H. Wiseman, "The two Bell's theorems of John Bell" *J. Phys. A: Math. Theor.* **47** 424001 (31pp).

[15] J.S. Bell, "The theory of local beables," *Epistemological Letters* March 1976; reprinted in *Speakable and Unspeakable*, *op cit.*

[16] J.S. Bell, "Free variables and local causality," *Epistemological Letters* Feb. 1977; reprinted in *Speakable and Unspeakable*, *op cit.*

[17] J.S. Bell, "EPR correlations and EPW distributions," *New Techniques and Ideas in Quantum Measurement Theory*, New York Academy of Sciences, 1986; reprinted in *Speakable and Unspeakable*, *op cit.*

[18] J.S. Bell, "La nouvelle cuisine," *Between Science and Technology*, A. Sarlemihn and P. Kroes, eds., Elsevier Science Publishers, 1990; reprinted in *Speakable and Unspeakable*, *op cit.*

[19] Tim Maudlin, "What Bell Did" *J. Phys. A: Math. Theor.* **47** (2014) 424010

[20] J.S. Bell, "Quantum mechanics for cosmologists," in *Quantum Gravity* **2**, C. Isham, R. Penrose, and D. Sciama, eds., Clarendon Press, Oxford (1981) 611-37. Reprinted in *Speakable and Unspeakable in Quantum Mechanics*, Cambridge, 2nd edition, 2004.

[21] J.S. Bell, "Against 'Measurement"' reprinted in *Speakable and Unspeakable in Quantum Mechanics*, Cambridge, 2nd edition, 2004.

[22] T. Maudlin, "Space-time in the quantum world" in *Bohmian Mechanics and Quantum Theory: An Appraisal*, J.T. Cushing, A. Fine and S. Goldstein, eds., Kluwer, 1996

[23] http://www.worldsciencefestival.com/2014/05/measure-for-measure-quantum-physics-and-reality/

[24] C. Fuchs, "A Formalism and an Ontology for QBism," draft grant application, shared via private communication Aug. 25, 2014.

[25] C. Fuchs, "QBism, the Perimeter of Quantum Bayesianism", arxiv:1003.5209

[26] C. Fuchs, N.D. Mermin, and R. Schack, "An Introduc-

tion to QBism with an Application to the Locality of Quantum Mechanics", *American Journal of Physics* **82** (8) August 2014, pp. 749-754

[27] N.D. Mermin, "Putting the Scientist into the Science", talk at *Quantum [Un]speakables II: 50 years of Bell's theorem*, online at `https://phaidra.univie.ac.at/detail_object/o:360625`

[28] C. Fuchs, "My Struggles with a Block Universe," Blake Stacey, ed., arxiv:1405.2390

[29] C. Fuchs and R. Schack, "Quantum-Bayesian Coherence: The No-Nonsense Version" *Rev. Mod. Phys.* **85**, 1693-1715 (2013)

[30] T. Norsen, "Einstein's Boxes" *American Journal of Physics* **73** (2) February 2005, pp. 164-176

[31] N.D. Mermin, "QBism puts the scientist back into science" *Nature* Vol. 507, Issue 7493 (26 March, 2014)

[32] C. Fuchs (with M. Schlosshauer and B. Stacey), "My Struggles with the Block Universe", arxiv:1405.2390

[33] J.S. Bell, "Toward an Exact Quantum Mechanics" in *Themes in Contemporary Physics II: Essays in honor of Julian Schwinger's 70th birthday*, S. Deser and R.J. Finkelstein, eds., World Scientific, 1989.

[34] R. Tumulka, "A relativistic version of the Ghirardi-Rimini-Weber model," *J. Stat. Phys* **125** (2006) p. 825

[35] D. Wallace and C. Timpson, "Quantum Mechanics on Spacetime I: Spacetime State Realism" *British Journal for the Philosophy of Science* **61** (4) pp. 697-727 (2010)

[36] D. Wallace, "Decoherence and Ontology" in *Many Worlds? Everett, Quantum Theory, and & Reality*, S. Saunders, J. Barrett, A. Kent and D. Wallace, eds., Oxford, 2010

[37] D. Wallace, *The Emergent Multiverse*, Oxford, 2012

[38] D. Bedingham, D. Dürr, G.C. Ghirardi, S. Goldstein, R. Tumulka, and N. Zanghí, "Matter Density and Relativistic Models of Wave Function Collapse," *J. Stat. Phys.* **154** (2014) pp. 623-631

[39] T. Maudlin, "What Bell proved: A reply to Blaylock" *American Journal of Physics* **78** (1) January 2010 pp. 121-125.

[40] V. Allori, S. Goldstein, R. Tumulka, and N. Zanghí, "Many-Worlds and Schrödinger's First Quantum Theory" *British Journal for the Philosophy of Science* **62**(1) 1-27 (2011)