Quantum Mechanics and Ordinary Language: The Fuzzy Link

Peter J. Lewis†‡

It is widely acknowledged that the link between quantum language and ordinary language must be "fuzzier" than the traditional eigenstate-eigenvalue link. In the context of spontaneous-collapse theories, Albert and Loewer (1996) argue that the form of this fuzzy link is a matter of convention, and can be freely chosen to minimize anomalies for those theories. I defend the position that the form of the link is empirical, and could be such as to render collapse theories idle. This means that defenders of spontaneous-collapse theories must gamble that the actual form of the link renders such theories tenable.

1. Interpreting Quantum States. On a realist construal of quantum mechanics, the quantum state determines the truth-values of claims about macroscopic objects. But what precisely is the connection between quantum mechanical descriptions and ordinary language descriptions of objects? The connection suggested by a literal-minded reading of the quantum formalism is the so-called *eigenstate-eigenvalue link,* which says that an object has a given property if and only if its state is an eigenstate of the corresponding operator. But this is widely acknowledged to be at best an idealization, since the states of macroscopic objects will never be *precisely* eigenstates of the properties we tend to attribute to them, such as position.

However, there are various ways of formulating quantum mechanics such that the states of macroscopic objects are generally *close* to eigenstates of the properties we attribute to them; spontaneous-collapse theories in the GRW tradition (Ghirardi, Rimini, and Weber 1986) are paradigm examples.[1] The existence of such formulations suggests a connection be-

---

†To contact the author, please write: University of Miami, P. O. Box 248054, Coral Gables, FL 33124-4670; e-mail: plewis@miami.edu.

‡I would like to thank Jeff Barrett, Jeremy Butterfield, P. D. Magnus, and Bradley Monton for helpful comments on an earlier version of this paper.

[1] It is worth noting that other formulations of quantum mechanics will need to loosen the link between eigenstates and ordinary language too. The results of this paper are not limited

Philosophy of Science, 70 (December 2003) pp. 1437–1446. 0031-8248/2003/7005-0048$10.00

tween quantum language and ordinary language that incorporates a measure of closeness to the relevant eigenstate. The most straightforward such measure, formulated by Albert and Loewer (1996), has been dubbed the *fuzzy link* (Clifton and Monton 1999).[2]

This paper concerns the status of interpretation rules like the fuzzy link. Albert and Loewer have argued that the precise form of the fuzzy link is not an empirical matter, but is more like a linguistic convention (1996, 91). This is handy, as it means that one can freely choose whatever interpretation rule provides the best fit between our quantum mechanical descriptions and our ordinary language descriptions of systems. But I will make a case here that Albert and Loewer's argument is a red herring, and that the form of the interpretation rule *is* an empirical matter. This creates something of a problem for the interpretation of quantum mechanics generally, since the choice of a spontaneous-collapse theory as a solution to the measurement problem amounts to a gamble that the correct interpretation rule for quantum mechanical states will provide an acceptable fit between wave-function language and ordinary language.

2. The Fuzzy Link. Before getting to the details of Albert and Loewer's argument, I first need to outline more precisely the connection between quantum mechanical language and ordinary language contained in the fuzzy link. Consider first the link between the physical descriptions of systems and ordinary language presupposed by *classical* physics. For a system consisting of a marble and a box, there are some configurations of particles for which the marble is in the box, some configurations of particles for which the marble is outside the box, and some configurations of particles to which neither description applies. The lines between these three sets of configurations are *vague,* in that there does not appear to be any principled way to draw precise distinctions between them.

To this classical vagueness, the fuzzy link adds an extra layer of quantum vagueness. The quantum mechanical wave function lives in configuration space; each classical particle configuration corresponds to a point in this space. So the (vaguely defined) set of particle configurations for which the marble classically counts as being in the box corresponds to a (vaguely defined) region of the configuration space—call it the "in" region. Similarly, the set of particle configurations for which the marble classically counts as being outside the box corresponds to a separate region of configuration space—call it the "out" region. An appropriate operator for

---

to spontaneous-collapse formulations, but for reasons of space this is the only case I discuss here.

[2] Ghirardi, Grassi, and Benatti (1995) suggest a somewhat different measure, but the arguments below do not rely on any feature of the fuzzy link not shared by their measure.

marble location could have any wave function contained entirely within the "in" region as an eigenstate with eigenvalue 1, and any wave function contained entirely within the "out" region as an eigenstate with eigenvalue 0.

Now according to the dynamics of a spontaneous-collapse theory, the wave function will never be entirely contained in either region, but will be spread out over them both (and over the rest of configuration space). However, it is perfectly possible for the wave function to be *almost* entirely contained in the "in" region, in the sense that the integral of the squared wave-function amplitude over the region is close to 1. In that case, the particle counts as being in the box according to the fuzzy link. Similarly, it is perfectly possible for the wave function to be almost entirely contained in the "out" region, in which case the particle counts as being outside the box according to the fuzzy link. The additional layer of vagueness comes about because there does not seem to be any principled way to specify precisely how much of the wave function needs to be in a given region in order for the marble to count as having the corresponding location.

3. The Argument of Albert and Loewer. Given this description of the fuzzy link, let us turn to Albert and Loewer's argument that the precise form of the link is conventional rather than empirical. Suppose we consider a precise version of the fuzzy link that specifies that the marble counts as being in the box if and only if the proportion of the squared amplitude in the relevant region is greater than or equal to $1-p$. Albert and Loewer argue that the value of p is not an empirical matter (1996, 90).

The argument goes as follows. Suppose there were an experiment that could determine whether or not p is smaller than 0.2. Such an experiment must reliably produce one of two distinct observable outcomes, say a check mark in one box if the value of p is less than 0.2 and a check mark in another box if p is greater than 0.2. But note that p does not appear in the dynamical laws, so the evolution of the wave function does not depend on the value of p. Hence the wave function at the end of the experiment will be the same whatever the value of p. So the only way in which the location of the check mark can depend on the value of p is if the application of the fuzzy link with p less than 0.2 to the final wave function entails that the check mark is in the first box, and the application of the fuzzy link with p greater than 0.2 to the same wave function entails that the check mark is in the second box. But note that the fuzzy link with p less than 0.2 is a more stringent condition than the fuzzy link with p greater than 0.2, so for any given wave function, whatever the former condition entails about the locations of macroscopic objects, the latter condition entails as well. So if the check mark counts as being in the first box with p set below 0.2, it must also count as being in the first box with p set above 0.2. So there is no way the location of the check mark could possibly depend on the value of p, so the value of p is not an empirical matter.

This is convenient, because it means we can freely choose the value of p in order to minimize a number of anomalies that arise when we attempt to connect the language of spontaneous-collapse theories to ordinary language. If p is greater than 1/2, it is possible, according to the fuzzy link, for an object to have two distinct locations at once; to avoid this anomaly, we can stipulate that p is less than 1/2. If p is too close to zero, then a spontaneous-collapse dynamics will not ensure that there are facts about the locations of objects like marbles; so we can avoid choosing p too close to zero. Unless p is precisely zero, conjunction introduction fails for claims about ordinary objects (Lewis 1997); so we can choose p sufficiently small that the anomaly arises only for sets of objects that are impractically large.

The upshot of Albert and Loewer's argument is that the precise form of the fuzzy link is not an empirical matter, and so we are free to choose it such that the language of quantum mechanics and ordinary language fit together cleanly and simply in all practical contexts. Clifton and Monton endorse this conception of the fuzzy link: "The fuzzy link, for some particular value of p, would then have something of the status of a postulate that (to echo Reichenbach above) 'is neither true nor false, but a rule which we use to simplify our language'" (1999, 716).

4. A Classical Analogue. I think the above argument fails to establish that the form of the fuzzy link is not empirical. The problem, however, is subtle, and I will approach it rather indirectly.

The first thing to note is that there is nothing particularly quantum mechanical about the argument as it stands. One could give a parallel argument that the precise form of the interpretation rule linking the language of *classical* mechanics and ordinary language is also not an empirical matter. Now it is hard to state any precise version of the link between classical particle configurations and ordinary language, so let me simplify the marble and box example somewhat. Consider a particle configuration in which all the particles are in the box, and which is otherwise such that the marble counts as being in the box, and another in which all the particles are outside the box, and which is otherwise such that the marble counts as being outside the box. Consider a process that transforms one configuration to the other by moving one particle at a time, thus defining a linear scale with the "in" configuration at one end and the "out" configuration at the other. Clearly the cutoff point between those configurations for which the marble counts as being in the box and those for which it does not is vague.

Suppose now that we specify a precise version of the classical analogue of the fuzzy link, which states that the marble counts as being in the box if and only if the proportion of particles remaining in the box is greater than or equal to $1 - p$. We can then apply Albert and Loewer's argument to show that the value of p is not an empirical matter. For the value of p to be empirical, there must be some experiment that reliably results in (for example) a check mark in one box if p is less than 0.2, but in another box if p is greater than 0.2. But p does not appear in the classical particle dynamics, so the particle configuration of the apparatus at the end of the experiment cannot depend on the value of p. So if the location of the check mark depends on the value of p, it must be because the final particle configuration is one such that there is a check mark in one box according to the classical fuzzy link with p less than 0.2, but in the other box according to the classical fuzzy link with p greater than 0.2. But as before, the version with p less than 0.2 is a more stringent condition than the version with p greater than 0.2, so whatever the former says, the latter says too. So if the final state is one in which the check mark is in the first box for p less than 0.2, then it is also a state in which the check mark is in the first box for p greater than 0.2. The location of the check mark cannot depend on the value of p, so the precise form of the classical version of the fuzzy link is not an empirical matter either.

But this is odd. Surely it's easy (in principle) to construct an empirical test that at least constrains the value of p. You simply have to watch the marble as its particle configuration changes from the initial "in" configuration to the final "out" configuration. To begin with, the particle will clearly be in the box, eventually it will clearly be outside the box, and in the middle, there will be some strange states that you don't know how to classify. At the end of this experiment, it seems, you could say, based on your own experience, that the value of p must lie somewhere within the range corresponding to the states you do not know how to classify, and not outside that range. But if the classical version of Albert and Loewer's argument is correct, no such empirical constraint on the value of p ought to be possible.

Let me put it another way. Suppose (as seems obvious) that you can at least empirically distinguish some particle configurations for which the marble is clearly in the box. This places empirical constraints on the value of p. But if the above argument is correct, then there are no empirical constraints on the value of p. So it follows that you cannot tell whether any particle configuration is one in which the marble counts as being in the box. But that is absurd; of course you can tell whether there is a marble in the box.

So what is wrong with the classical analogue of Albert and Loewer's argument? What is wrong, I think, is the description of the kind of experimental result that would be necessary in order to fix the value of p empirically. I followed Albert and Loewer in assuming that an experimental determination requires two check boxes; a check mark in one box indicates that p is less than 0.2 and a check mark in the other indicates that p is greater than 0.2. But why not make do with one check box? If there is a check mark in the box, that indicates that p is greater than 0.2, and if not, that indicates that p is less than 0.2. Given such a set-up, the classical argument I sketched above no longer works. It is perfectly possible for the final particle configuration in the experiment to be such that the check mark counts as appearing in the box according to the less stringent version of the fuzzy link (with p greater than 0.2) but does not count as appearing in the box according to the more stringent version (with p less than 0.2). So it is perfectly possible to empirically constrain the value of p.

In fact, when I described the experiment in which you simply watch a marble while its particle configuration changes from an "in" configuration to an "out" configuration, what I was describing was essentially a test of the above kind. Substituting for the single check box in this case is you, or more precisely, your brain. During the experiment, your brain-state can be described (classically) by some particle configuration. This particle configuration may be such that on a less stringent version of the classical fuzzy link, it counts as one in which you see a marble in the box, but on a more stringent version, it does not count as one in which you see a marble in the box. And what it counts as is something you have direct access to; you know whether you're seeing a marble or not, and this constrains the value of p.

5. Observation in Spontaneous-Collapse Theories. In the classical case, then, your experience directly constrains the precise form of the classical fuzzy link. Does the same move work for the original quantum mechanical version of the fuzzy link? That is, can I counter Albert and Loewer's argument by describing a quantum mechanical version of the direct observation experiment? I think the answer is "yes," but some difficulties need to be dealt with.

Superficially, it's easy to describe a similar experiment for the quantum case. You watch a marble whose wave function is gradually altered from one with 100 percent of its squared amplitude in the "in" region of configuration space to one with 100 percent of its squared amplitude in the "out" region. To begin with, the marble is clearly in the box, and at the end it is clearly outside the box, and in the middle you don't know what to say. Where these transitions take place constrains the value of p.

But there are serious difficulties with this thought experiment. According to any spontaneous-collapse dynamics, almost all of these marble states will be highly unstable. In particular, the states away from the two extremes will all spontaneously and rapidly evolve to states much closer to one of the extremes. Even if you could prepare, say, a state in which 70 percent of the squared wave function amplitude is in one region and 30 percent is in the other, it will rapidly collapse to a state in which close to 100 percent of the squared amplitude is in one or other of the regions. Further collapses will occur within your eyes and brain when you observe the marble, intensifying the effect. So it looks like the direct-observation experiment is not practically possible in the case of spontaneous-collapse theories.

While I don't think this practical difficulty counts against the claim that the value of p is empirically constrained, it does require me to explain the nature of the empirical constraint more carefully. Note first that what most directly determines your experience is the state of your brain; your experience that there is a marble in the box is most directly explained in terms of a particular brain state. Since the spontaneous-collapse program is designed to ensure that locations can be ascribed to medium-sized lumps of matter, it seems to be a precondition of the program that experience supervenes on the locations of many-particle subsystems of the brain.[3] A version of the fuzzy link will apply to these brain states; you have a particular experience (say, seeing a marble in a box) if and only if a certain proportion of squared wave-function amplitude (say, more than $1-p$) is in the relevant region of configuration space.

Now consider three states: a "sees-in" eigenstate in which all the squared wave-function amplitude for your brain is in a "seeing a marble in the box" region of its configuration space, a "sees-out" eigenstate in which all the squared amplitude is in a "seeing the marble outside the box" region, and a superposition state in which most of the squared amplitude is in the former region but a little is in the latter region. In order for the superposition state to be close enough to the "sees-in" eigenstate to count as a state in which you see the marble in the box, it must be subjectively indistinguishable from the "sees-in" eigenstate but subjectively distinct from the "sees-out" eigenstate. Whether or not this is in fact the case is clearly an empirical matter. And whether or not it is the case constrains the value of p; the value of p must be such that it lumps together subjectively indistinguishable brain states.

Typically, when you observe a marble, the collapse dynamics will result in a superposition state which is either very close to a "sees-in" eigenstate or very close to a "sees-out" eigenstate. The empirical question is whether these states are close enough to the relevant experience eigenstates to be subjectively indistinguishable from them, and my experience of marbles suggests that in fact they are close enough. So experience provides a lower bound for p; it cannot be so small as to entail that the typical post-collapse

---

[3] Albert (1992, 104–106) makes this requirement explicit; Aicardi et al. (1991) argue that it is actually satisfied by human brains.

states do not count as close enough to the relevant experience eigenstates according to the fuzzy link. Furthermore, given this lower bound, the states of the marbles themselves will generally count as close enough to the relevant location eigenstates according to the fuzzy link. Since I have no experiences of states far from the eigenstates, I have no experiences that could provide an upper bound for *p,* but this does not affect my central claim, contra Albert and Loewer, that the value of *p* is empirically constrained.

6. Objects in Two Places at Once. So far this may not seem terribly interesting. If one takes the view that the precise form of the fuzzy link is conventional, as Albert and Loewer do, then one can *choose* the value of *p* so that the fuzzy link ascribes locations to macroscopic objects. On the other hand, if one takes the view that the precise form of the fuzzy link is empirical, as I maintain, then one's *experience* provides a lower bound for *p,* so again the fuzzy link ascribes locations to macroscopic objects. Either way, the value of *p* is such that we can ascribe locations to objects, and so the reliance of spontaneous-collapse theories on a measure like the fuzzy link is unproblematic.

However, I think there *is* a potential problem with the use of the fuzzy link, one that arises at the other end of the range of possible *p*-values. Although our experience constrains *p* to be large enough that objects have locations, it does not constrain *p* to be small enough that objects generally have only a *single* location—at least, not in the straightforward way one might think.

Albert and Loewer simply stipulate that *p* is less than 1/2, which automatically prevents an object from having two locations. But if the value of *p* is empirical, whether *p* is less than 1/2 cannot be simply a matter of stipulation. Still, one might expect from the foregoing that it would be easy to argue *empirically* that *p* is less than 1/2. After all, don't we know from *experience* that everyday objects like marbles aren't found in more than one place at a time?

But unfortunately it is not that simple. Suppose *p* is close to 1, so that according to the fuzzy link, an object counts as having a particular location if even a small proportion of the squared wave function amplitude is in the corresponding region of configuration space. Then a state in which most of the squared wave function amplitude is in the "in" region but a small amount of it is in the "out" region may count as a state in which the marble is in the box and *simultaneously* count as a state in which the marble is outside the box. And if you observe the marble, your brain ends up in a state in which most of the squared wave function amplitude is in the "seeing a marble in the box" region of its configuration space, but a small amount of it is in the "seeing a marble outside the box" region. This state may count as one in which you see the marble in the box and *simultaneously* as one in which you see the marble outside the box. However, since the two terms in the state don't interact with each other, you won't be *aware* of the fact that you have two conflicting experiences. Your experiences from this point on form two distinct streams, neither of which has access to the other.[4]

So for all you know, the value of p could be large enough that macroscopic objects have simultaneous incompatible positions, and hence large enough that your experience regularly "splits" into incompatible streams. But if that were the case, then a spontaneous-collapse account of quantum mechanics would completely beside the point, because it would make no difference whether the states of systems are generally close to eigenstates for the properties we ascribe to them. If p is sufficiently large, it makes no difference whether 50 percent or 99.9 percent of the squared wave function amplitude is in the "in" region; in either case, the marble will count as being both in the box and outside the box. And similarly, it makes no any difference whether 50 percent or 99.9 percent of the squared wave function amplitude is in a "seeing the marble in the box" region; in either case, you have two incompatible experiences. So if p is large enough (and for all we know it could be), then any spontaneous-collapse mechanism is idle.[5]

7. The Strange Status of the Fuzzy Link. The above argument might seem to count against my claim that the value of p is an empirical matter. However, even though the upper bound for p is not *straightforwardly* constrained by our experience, there is still an important sense in which it is an empirical matter. There are experiments one could perform that would provide strong evidence about the value of the upper bound. Several authors have shown that if something like the above "splitting-experience" view of quantum mechanics were true, it would follow that people are immortal (Price 1996, 221; Lewis 2000). For example, if you shoot yourself in the head, there is a small term in the wave function in which the bullet spontaneously dissociates before it hits you. But if the value of p is large enough, this small term is enough to ensure the continuation of a stream of experience for you, and this in turn is enough to ensure your survival. And the same goes for any potential threat to your life.

This provides a possible empirical test for the value of p. If you shoot yourself in the head and the bullet spontaneously dissociates before it hits you, you can be pretty sure that the value of p is large enough to render

---

[4] See Lockwood (1989). Lockwood (quite plausibly) interprets each stream of experience as a separate *person*.

[5] Cordero (1999) raises a similar criticism of spontaneous-collapse theories. Unlike Cordero, I do not think this problem is inevitable, as it only arises given a particular form of the fuzzy link.

spontaneous-collapse theories idle; otherwise you would have witnessed a statistical miracle. Of course, you would be crazy to perform such an experiment. And even if you did, the result would only be known to *you;* any observers would almost certainly observe your death, but this would tell them nothing about the value of *p*.[6] This puts us in a curious situation with regard to the fuzzy link. On the one hand, the precise form of the link between wave function language and ordinary language is an empirical matter. On the other hand, there is no reasonable way to *find out* about this empirical fact. And to make matters worse, the tenability of spontaneous-collapse theories *depends* on this fact.

Albert and Loewer argue that spontaneous-collapse theories are unproblematic because we can simply stipulate an appropriate link between wave-function language and ordinary language. But if the foregoing is correct, the form of the link isn't a matter of stipulation. Unbeknownst to us, the actual form of the link may rule out spontaneous-collapse theories altogether, and point to a no-collapse theory of the "splitting experience" variety instead. Furthermore, this difficulty would appear to afflict any attempt to make the wave function alone the truthmaker for our ordinary claims about macroscopic objects. Hence anyone proposing such a theory is essentially gambling that the form of the link between quantum language and ordinary language will turn out to be favorable to their project.

REFERENCES

Aicardi, F., A. Borsellino, G. C. Ghirardi, and R. Grassi (1991), "Dynamical Models for State-Vector Reduction: Do They Ensure That Measurements Have Outcomes?", *Foundations of Physics Letters* 4: 109–128.

Albert, David Z. (1992), *Quantum Mechanics and Experience*. Cambridge: Harvard University Press.

Albert, David Z. and Barry Loewer (1996), "Tails of Schrödinger's Cat", in R. Clifton (ed.), *Perspectives on Quantum Reality*. Dordrecht: Kluwer, 81–92.

Clifton, Rob, and Bradley Monton (1999), "Losing Your Marbles in Wave-Function Collapse Theories", *British Journal for the Philosophy of Science* 50: 697–717.

Cordero, Alberto (1999), "Are GRW Tails as Bad as They Say?", *Philosophy of Science* 66: S59–S71.

Ghirardi, G. C., R. Grassi, and F. Benatti (1995), "Describing the Macroscopic World: Closing the Circle within the Dynamical Reduction Program", *Foundations of Physics* 25: 5–38.

Ghirardi, G. C., A. Rimini, and T. Weber (1986), "Unified Dynamics for Microscopic and Macroscopic Systems", *Physical Review* D34: 470–491.

Lewis, Peter J. (1997), "Quantum Mechanics, Orthogonality and Counting", *British Journal for the Philosophy of Science* 48: 313–328.

——— (2000) "What Is It Like to Be Schrödinger's Cat?", *Analysis* 60: 22–29.

Lockwood, Michael (1989), *Mind, Brain and the Quantum*. Oxford: Blackwell.

Price, Huw (1996), *Time's Arrow and Archimedes' Point*. Oxford: Oxford University Press.

---

[6] See Lewis (2000, 27). Briefly, the reason is that when each observer's experience splits, the usual quantum mechanical probabilities apply, and so the term in which you survive has extremely low probability. For you, however, only one stream of experience remains after the experiment, namely the one in which you survive.