David Papineau

David Lewis and Schrödinger's Cat

Abstract

In 'How Many Lives has Schrödinger's Cat?' David Lewis argues that the Everettian no-collapse interpretation of quantum mechanics is in a tangle when it comes to probabilities. This paper aims to show that the difficulties that Lewis raises are insubstantial. The Everettian metaphysics contains a coherent account of probability. Indeed it accounts for probability rather better than orthodox metaphysics does.

1. Introduction

On 27 June 2001, not four months before his untimely death, David Lewis delivered the third Jack Smart Lecture at the Australian National University. His title was 'How Many Lives has Schrödinger's Cat?'[1] and he spoke on a topic which is absent from his published writings, the interpretation of quantum mechanics. More specifically, he discussed the 'no-collapse' interpretation of quantum mechanics pioneered by Hugh Everett III [1957]. Lewis allowed that this interpretation offers initial theoretical attractions, but also argued that it suffers from irremediable flaws.

If you have heard of Everett because of his association with the 'many-worlds' interpretation of quantum mechanics popularised by Bryce Dewitt [1970, 1972], you might suppose that there is some affinity between the no-collapse interpretation and Lewis's philosophical realism about possible worlds. But this would be a mistake. While Everett's interpretation does add extra 'branches' to the reality recognized by common sense, these additions fall far short of Lewis's multiplication of worlds. For a start, the extra 'branches' that Everett adds to reality all lie within the actual world that evolves from the actual initial conditions in line with the actual laws of physics— these branches by no means include all possibilities. Moreover, Everett's branches are best conceived, not as sunderings of the whole universe, but rather as entities that spread out causally at finite speeds, 'like ripples on a pond', as Lewis puts it. For example, in the Schrödinger's Cat experiment, <u>first</u> the photon branches into a deflected and undeflected version when it passes through the half-slivered mirror; <u>then</u> the detector branches into a triggered and untriggered state when it interacts with the photon; <u>then</u> the poison bottle branches into a smashed bottle and an unsmashed bottle under the influence of the detector; and so on, culminating in the cat branching into a live and dead cat, and the human observer branching into a self who sees a live cat and a self who sees a dead cat.[2]

---

[1] The lecture is reprinted on pp. 00-00 of this volume.

[2] 'Many-worlds' is thus not an apt name for the optimal Everettian view, as Lewis observes (p. 00). Following Lewis's lead, I shall stick to the simple 'no-collapse interpretation'. (I used to favour the terminology of 'many minds' as a way of conveying the local nature of Everettian splitting [Papineau 1995; 1997] but I now think that this suggests an overly subjective reading of Everett.)

It is precisely this causal proliferation of branches that makes the no-collapse interpretation so theoretically attractive. Basic quantum mechanics leaves us with no alternative but to allow that <u>microscopic</u> entities can be in superpositions of different observable values—for example, in the double-slit experiment we need to allow that the electron's state prior to observation contains branches corresponding to its passage through each of the slits, otherwise we will not be able to explain the observed interference effects. However, this means that orthodox quantum mechanics has great difficulty in accounting for the apparent definiteness of the macroscopic world. For the central quantum mechanical law of motion, Schrödinger's equation, indicates that macroscopic systems which interact with other superposed systems will themselves enter into superpositions, just as in the Everettian reading of the Schrödinger's Cat experiment outlined above. So, in order for orthodox interpretations of quantum mechanics to stop this disturbing proliferation of macroscopic branches, they must add something to Schrödinger's equation—and they add 'collapses'. At some point reality 'collapses' unpredictably into just one element of the proliferating network of superpositions. The trouble facing such orthodox collapse interpretations, however, is that all explicit theories about 'collapses' seem arbitrary and ad hoc (not to mention their inconsistency with special relativity and with the conservation of energy). This is where Everett has an advantage. It simply embraces the deterministic Schrödinger evolution of causally proliferating superpositions, and denies that reality ever collapses into just one branch. By thus rejecting collapses, Everett thus promises to remove 'a gratuitous blotch on an otherwise elegant theory', as Lewis puts it.

The challenge facing the Everett interpretation is to explain how we human beings fit into this strange world of proliferating branches. This is not just a matter of anthropocentric self-absorption. For unless we can show how the no-collapse interpretation saves the appearances—that is, how it predicts what <u>we</u> experience—we will have no reason to believe it in the first place.

Following Lewis, we can divide this challenge into two parts. The first is to explain why our experience should always present the world as <u>definite</u>, when in fact it is in a superposition. The second is to explain in what sense the no-collapse interpretation can predict that some braches are more <u>probable</u> than others, given that it also says that all branches will definitely occur.

On the first issue, Lewis concedes there is no difficulty. He offers the model of a duplicating beamer-upper that gives you two successors, one on the starship <u>Enterprise</u> and another on the starship <u>Ptomekin</u>. These successors will each have their own experiences, and before you are beamed up you can anticipate becoming both successors (though not becoming one successor who has both sets of experiences). The same will occur when you 'branch' after interacting with a quantum system that is in a superposition of two definite observable values. After the interaction, you will have two successors, one of whom observes one value and another of whom observes the other value—and before the interaction you can anticipate becoming both successors (though not becoming one successor who has both sets of experiences).

It is the second issue, to do with probabilities, that Lewis takes to pose problems for Everett. The main body of Lewis's paper, after his elegant and illuminating explanation of the above issues, is concerned to show that Everettians are in a tangle

when it comes to probabilities. Not that Lewis aims to present a knock-down argument against the no-collapse view. Rather he makes two points. First, he maintains that Everettians have no good way of justifying the 'intensity rule' which is their only alternative to orthodox probabilistic thinking. Second, he argues that Everettians are forced to discount the possibility of death in life-or-death situations, and that this has unpalatable consequences.

As I said, Lewis does not present any of this as conclusively discrediting the no-collapse interpretation. But he clearly takes it to tip the balance against Everett. Thus he mentions in section 2 that his own marked preference is for a collapse theory along the lines of the 'GRW hypothesis' (while simultaneously admitting that his readiness to countenance collapses may be due to his having a scientific background in chemistry rather than mathematical physics) [Ghirardi, Rimini and Weber 1986].

I what follows I shall try to show that the difficulties that Lewis raises for Everett are insubstantial. The picture of reality offered by the no-collapse interpretation is certainly weird and wonderful. But it is quite cogent, and in particular it contains a coherent account of probability. If worries about probability are the main reason for dismissing Everett and continuing to put up with collapses, then maybe we should think again. If you ask me, even chemists have reason to take Everett seriously.

The rest of this paper contains two sections. The next section discusses the Everettian 'intensity rule'. The final section considers whether Everett has unpalatable implications for life-and-death situations.

2. The Intensity Rule

2a. Chances and Intensities

Orthodox collapse interpretations of quantum mechanics ascribe <u>chances</u> to the various possible outcomes of a collapse, corresponding to the pre-collapse squared amplitudes of the branches that will yield these outcomes. Only one of the alternative outcomes will become real. The chance of an outcome can be thought of as signifying its current standing in the competition to become real. Sometimes outcomes with low chances will become real, but the odds are in favour of outcomes with high chances.

The no-collapse interpretation has no room for chances, so conceived. For it denies that only one of the alternative outcomes will occur. From the Everettian point of view there is no competition to become real. All the alternatives are determined to happen. This makes it hard to see how any of them can have a chance different from one.

Lewis points out that without chances the no-collapse interpretation is in danger of sawing off the branch that it is sitting on. After all, the only reason for believing quantum mechanics in the first place, as with any theory, is that the world we observe matches the world that the theory advises us to expect. But the expectations advised by an indeterministic theory like quantum mechanics derive from the chances it implies. Any indeterministic theory tells us to expect outcomes in proportion to its chances, and is then confirmed to the extent that it urges high expectations for what is

actually observed. However, if the no-collapse interpretation eschews chances, on the grounds that all alternative outcomes will occur, then it would seem unable to advise us to expect different outcomes to different degrees, and so unable to maintain that quantum mechanics is confirmed by what is actually observed.

Lewis allows that Everettians might respond by adopting an 'intensity rule' to govern their expectations. Suppose that Everettians think of their non-collapsing branches as having differering 'intensities', corresponding to the squared amplitudes of those branches. Then they can proportion their expectations about the future directly to these intensities, even in the absence of any chances. For example, in the case of Schrödinger's cat, if the squared amplitude of the branch where the cat is alive is 50%, and that of the branch where it is dead is 50%, then you should expect to see the cat alive to degree 50%, and expect to see it dead to degree 50%. The intensity rule thus offers Everettians an alternative route to the confirmation of quantum mechanics. If you proportion your expectations directly to the squared amplitudes, in line with the intensity rule, then once more you can regard quantum mechanics as confirmed if the outcomes you observe are the ones that you expected.[3]

Lewis's central worry about the intensity rule is that there is no good way to justify it. Everettians have no option but to accept it as a primitive truth about rationality.

Still, as Lewis himself immediately asks 'How bad is that?' After all, orthodox metaphysics itself adopts an analogous 'chance rule', advising minds to proportion their expectations to the chances. In particular, orthodox interpretations of quantum mechanics presume just this rule when they advise expectations about the outcomes of collapses. Yet, as Lewis admits, there is no good way of justifying the chance rule either.

This suggests that Everett is here on the same footing as orthodoxy. Everett needs the intensity rule. Orthodoxy needs the chance rule. Neither can be justified.

In fact I think the situation is even more favourable to the no-collapse view than this. When it comes to the justification of rules telling us how to proportion expectations to the objective facts, not only is the no-collapse view no worse off that orthodoxy (as I shall argue in section 2b)—it is positively better off. Indeed it is better off twice over. First, orthodoxy faces an extra issue of justification, which does not arise on the no-collapse perspective (section 2c below). Second, the no-collapse perspective promises to ground the intensity rule in physical symmetries, in ways not obviously available to orthodoxy (section 2d).

---

[3] The no-collapse interpretation implies that, in addition to those (high intensity) later selves who observe frequencies that confirm quantum mechanics, you will also have (low intensity) later selves who observe rogue frequencies that disconfirm quantum mechanics. This might seem worrying, but it is not clear that the no-collapse view has any more trouble with knowledge of quantum-mechanical amplitudes than does orthodoxy. On any account of statistical inference, there is always a danger of observing an improbable frequency in repeated trials. Even so, orthodoxy takes observed frequencies to be evidence for corresponding chances (and hence quantum mechanical amplitudes). Everettians can advise us to reason in just the standard way, modulo the substitution of intensities for chances: infer that the intensity (and hence quantum mechanical amplitude) is close to the observed frequency, and hope that you are not the victim of an unlucky sample.

By contrast, Lewis thinks that the intensity rule is more problematic than the chance rule. We can introduce each of the next three sections with arguments he offers, all alas frustratingly brief.

2b. Two Mysteries or One? Lewis's first argument is that the no-collapse interpretation will be stuck with two mysteries, where orthodoxy will have only one: Everett will need both the unjustifiable chance rule and the unjustifiable intensity rule, whereas orthodoxy needs only the chance rule. However, Lewis's argument for this claim is rather puzzling. He says the two rules 'are not at all the same sort of thing', observing that the chance rule is concerned with alternative possibilities, whereas the intensity rule deals with co-existing actualities. In consequence, says Lewis, the expectations governed by the intensity rule are not the same as the subjective probabilities dictated by the chance rule, nor are intensities the same as chances. Subjective probabilities and chances pertain to alternative possibilities, not co-existing actualities.

Maybe so, but I don't see why this shows that Everettians need two rules rather than one. Suppose I agree that the intensity and chance rules 'are not at all the same sort of thing'. This doesn't show that I have to adopt both, if I adopt the intensity rule. Why can't I adopt the intensity rule and drop the chance rule? And this is surely what any sensible Everettian will do, if persuaded by Lewis that the two rules are so very different. Instead of adopting the intensity rule as an addition to the chance rule, Everettians will simply replace the chance rule by the intensity rule.

Of course, this strategy will only make sense if the situations where the chance rule is needed are all quantum mechanical. For the intensity rule is explicitly tailored to quantum mechanical intensities, and so will only be able to replace the chance rule in set-ups where such intensities are available. If chances are sometimes present in non-quantum mechanical situations, then there won't be any intensities to govern Everettian expectations, and in these cases Everettians will need the chance rule in addition to the intensity rule. (For example, you might think that the toss of an ordinary coin involves chances but no quantum mechanics.)

Still, it seems perfectly plausible that all seriously chancy situations do in fact have a quantum-mechanical basis. In support of this, consider Lewis's own comments in section 5 of his paper, when he argues that all death-mechanisms are quantum-mechanical. At first sight it may seem that deaths due to poisoning, say, or shooting, or auto-immune disease, owe nothing to quantum mechanics. But Lewis argues convincingly that all biochemical and mechanical processes are subject to quantum mechanical unpredictability. In line with this, it seems open to Everettians to hold that any serious statistics displayed by such processes are reflections of underlying quantum mechanical intensities. Again, in previous work Lewis has argued that even ordinary coin tosses might be viewed as quantum-mechanical processes, which once more suggests that Everettians will be able to apply the intensity rule [1986: 118-19]. In general, then, it seems open to Everettians to argue that quantum-mechanical intensities will be available in all situations where orthodox thinking demands that expectations are tailored to chances.

So far I have assumed that someone who adopts the intensity rule as the sole principle governing expectations will therewith reject the chance rule. This is in line with

Lewis's claim that the two rules are 'not at all the same kind of thing'. However, this elimination of the chance rule is not the only option available to Everettians. An alternative is to argue that the intensity rule reduces the chance rule, by showing us more clearly the real nature of chances. According to this reductionist option, orthodox thinking is right to hold that expectations should reflect chances, but wrong to assume that chances are measures over competing possibilities: in truth, chances always were measures over co-existing branches of reality, and the intensity rule is nothing other than the chance rule stripped of outmoded metaphysics.

I do not take this choice between elimination and reduction to be a substantial issue. As with most such choices, there is no reason to suppose that the prior meaning of the crucial term ('chance') is definite enough to decide the issue [cf. Papineau 1996]. Is it part of the definition of 'chance' that it is a measure over competing possibilities (in which case Everettians must say there are no chances)? Or does the definition of 'chance' specify only that chance is that magnitude which expectations should reflect (in which case Everettians can equate chances with intensities)? Since nobody ever thought to stipulate an answer beforehand, it seems to me a matter of choice how we sharpen the term 'chance' now that the question has arisen.

For my money, it is a polemically better strategy for Everettians to say they are keeping chances, but thinking of them rather differently.[4] But I do not propose to press the point here. As I said, it is not a substantial issue. Either way, the Everettians will end up with one primitive rule governing expectations, and to this extent will be on a par with orthodoxy. Polemics aside, it doesn't matter whether they say it is a refinement of the old chance rule, or a replacement for it.

Still, it will be helpful to define our terms, and for the purposes of the subsequent discussion I am happy to understand 'chance' as Lewis does, tying it to an orthodox metaphysics of competing possibilities. I shall use 'intensity' for the Everettian measure over coexisting branches. When I want a term for objective single-case probability that is neutral between the two notions, I shall simply employ 'probability'.

2c. Caring for Future Selves Lewis thinks there is another reason why the lack of justification counts against the intensity rule more than the chance rule. 'Not only do we have no way to justify the intensity rule; we have a plausible way to justify a conflicting rule. All your future selves, on all your branches, are equally real, and equally yours. You will have experiences of all of them. Do they not deserve equal weight . . . regardless of their intensities?' (p. 00).

At first sight Lewis may seem to have a point here. The no-collapse intensity rule does not treat all future selves equitably, despite the full-blooded actuality of these multiple selves. However, it would be too quick to assume that this counts in favour of orthodoxy and against Everett. For orthodoxy can also be accused of inequitable

---

[4] In another context, Lewis himself makes a similar move. At the end of his 'Humean Supervenience Debugged' [1994], he observes that the kind of chances defended in that paper only imperfectly satisfy the role that defines 'chance'. He says: '. . . nothing perfectly occupies the role, so nothing perfectly deserves the name. But near enough is good enough. . . . an imperfect candidate may deserve the name quite well enough' (p. 489). (I offer reasons for thinking that Everettian intensities are indeed 'near enough' to ordinary chances in [Papineau 1995].)

bias in its treatment of future selves. What is more, orthodoxy's bias turns out to be in tension with its underlying metaphysics, in a way that the Everettian bias isn't. So in the end it is orthodoxy, not Everett, that has more trouble with the biased treatment of future selves.

Let me go more slowly. First consider Everett's treatment of future selves. By setting expectations equal to the intensity of the branches, Everett does indeed favour some successors over others. To see this, consider choices between actions that benefit your successors differentially. Insofar as you conform to the intensity rule, you will favour those actions that benefit your high-intensity successors over those that benefit your low-intensity successors.[5]

Still, does orthodoxy make uncertain choices any less discriminatory? According to orthodox metaphysics, in any chancy situation I will have a number of <u>possible</u> successors. Yet these successors do not weigh equally in orthodox choices either, since the chance rule analogously advises me to choose those actions that benefit my high-chance possible successors over those that benefit my low-chance possible successors.

You might feel inclined to respond that the two cases are different. On the orthodox view, only one of your possible successors will become actual, not all of them, and only <u>this</u> successor really matters. By contrast, on the no-collapse view all your possible successors will become actual, and so are all worth caring about. True enough. But this difference, far from helping orthodoxy, turns out to cause it extra problems. Think of it like this. Both orthodoxy and Everett advise acting with the probabilities, in the sense of favouring high-probability possible successors over low-probability ones. And both take this to be a primitive principle of rationality, since neither can justify it in terms of something more basic. But is it much odder for orthodoxy to take this principle as primitive than Everett, since orthodoxy doesn't <u>care</u> about all possible successors, in the way that Everett does. Orthodoxy thinks that only the one actual successor matters. So, even after primitively committing itself to acting with the probabilities, orthodoxy would seem to face an extra question, which doesn't arise for Everett, of explaining why this commitment is good for the sole actual successor who matters.

To see the point, note that we normally take the <u>aim</u> of an uncertain choice to be benefit to the sole actual successor. Yet in practice I choose that action that maximizes benefit over <u>all</u> possible successors weighted by their <u>chances</u>. This action won't necessarily benefit my actual successor (odds-on favourites can lose, and long shots can come home). So there is room to ask orthodoxy: why is it such a good idea to opt for the action that maximizes chance-weighted benefit over all possible successors, given that what I really want is benefit to my sole <u>actual</u> successor?

You might think that orthodoxy can justify maximizing chance-weighted benefit on the grounds that this ensures your actual successors will do well <u>in the long run</u>. But there is no guarantee that betting with the chances will win in the long run either. I can be unlucky in the long run as well as the short. Perhaps, if you are prepared to

---

[5] Doesn't this last claim presuppose that choices will maximize expected benefit? Yes, but this is no new assumption. It is already built into the notion of expectation: agents <u>expect</u> given outcomes to just the extent that those outcomes weigh in their choices.

dabble with frequency theories of chance, you might be inclined to argue that success will be guaranteed once I have exhausted all the cases that contribute to the frequencies that fix the chances. But this move only draws attention to a more fundamental objection to the long-run justification. The question at issue is: why is right to bet with the chances <u>now</u>, in this particular case? It is no answer to be told that so betting would comprise one component in a composite possible action that would guarantee eventual long-run success. What if I have no thought for the future, and am concerned only to make money on some bet <u>today</u>? Surely I still have just as much reason to bet with the chances as anybody else. [Cf. Pierce 1923: 69; Hacking 1965: 47; Putnam 1987: 80-4.]

Readers may feel I am pushing at an open door here. After all, isn't is agreed on all sides that there is no way of justifying the chance rule in terms of something else? But it is not the mere unjustifiability of the chance rule to which I am currently concerned to draw attention. Rather, I want to bring out how odd it is that this unjustifiable principle should advise us to choose one kind of action (that which maximizes benefit <u>over all my possible successors</u>), when what we really want is a different kind of action (that which maximizes benefit to my sole <u>actual</u> successor). It is one thing to adopt an unjustifiable principle advising actions of kind Φ. It is another thing to adopt such a principle when we don't care about Φ at all, but only about something else to which Φ has no non-question-begging connection.

I find orthodox thinking close to paradoxical on this point [cf. Papineau 2003a]. It is striking, however, that the threat of paradox is peculiar to orthodox metaphysics, and does not arise within the no-collapse framework. The puzzle for orthodoxy is to explain why I should maximize chance-weighted benefit over all possible successors, given that what I really want is benefit for my sole actual successor. On the Everettian view, though, this puzzle simply disappears. For the puzzle presupposes that only one of my possible successors will be actual ('my sole actual successor'). But if there are no collapses, then I have no sole actual successor—rather I will be succeeded by all possible successors, weighted by their intensities. So Everettians face no further puzzle, once they adopt the basic principle that I should act with the intensities (the intensity rule). If I have no unique future self, there is no need to explain why acting on the intensities is good for that unique future self. Rather, acting with the intensities is already benefiting all my actual future selves, in proportion to their intensities.

Speaking for myself, I find this a very persuasive argument in favour of the no-collapse interpretation. I have always found it very disturbing that there is no good way of justifying the chance rule. My disquiet wasn't just that the rule is unjustifiable. After all, justification has to stop somewhere. The real worry was that this seemed quite the wrong place for our spade to turn. Since the chance rule recommends that we perform actions with one feature (probable success) when we really desire another feature (actual success), it seems as if there ought to be some non-question-begging way of connecting the chance rule's recommendation with what we really desire. But there isn't.

From this perspective, it is orthodoxy that has two mysteries, where Everett only has one. Both start by advocating their respective rules (chance, intensity) as primitive. But Everettians can stop there, with one mystery. By contrast, orthodoxy faces the

further challenge of explaining why we should do things we don't care about in pursuit of things we do—a challenge it cannot answer.

2d. Constraints of Rationality

At the end of his section 3 Lewis says: 'But quantum mechanical intensity, unlike chance, is a recently discovered and theory-laden magnitude, unknown to all rational thinkers of the past and many rational thinkers of the present. It's not at all plausible that it might figure in any basic principle of rationality.' It is not immediately clear what to make of this. The chance rule and the intensity rule are both prescriptions to match expectations to some mind-independent feature of the world. There is no obvious reason why the better of these prescriptions should always have recommended itself to rational thinkers. Even so, it will be interesting to compare the two rules on this score. The intensity rule will turn out to fare much better than might be expected.

As a first step, it will be helpful to decompose both the chance and intensity rules into a priori and a posteriori components. This decomposition is familiar in connection with the chance rule. The chance rule can be split into (a) an a priori specification of the 'chance role', plus (b) an a posteriori account of the 'chancemakers' that fill this role in the actual world. The a priori component has been articulated by Lewis as the 'Principal Principle': in effect, this stipulates that chance is that quantity to which it is rational to proportion our expectations [Lewis 1980]. It is then a further question what mind-independent quantity actually fills this role. We can imagine a similar decomposition of the intensity rule into a priori and a posteriori components.

Let us first compare our rules in the a priori dimension. If there is to be an opposition here, we will need to build the metaphysical difference between chance and intensity into the stipulation of roles. (If the a priori chance role merely stipulates that chance is that quantity to which rational expectations should be proportioned, then Everettians can simply take the reductionist line that their intensities are the actual realizers of this role.) So let us accordingly build it into the stipulation of the chance role that chances are a measure over alternative possibilities, and correspondingly build it into the stipulation of the intensity role that intensities are measures over persistently coexisting actualities.[6][6]

Now, it is undeniable that the chance role, so understood, has been more familiar to rational thinkers through history than the corresponding intensity role. The intensity role assumes that probabilistic set-ups develop into coexisting actualities, and this is a recent innovation, prompted by the need to interpret quantum mechanics. The associated notion of intensity is therefore no more than a few decades old. However, it is not clear that chance itself does significantly better. If intensity is a few decades old, chance can at best boast a few centuries. For chance is also a 'recently discovered and theory-laden magnitude'. In The Emergence of Probability [1975] Ian Hacking shows that pre-seventeenth-century rational thought had no place for a

---

[6][6] This articulation of metaphysical presuppositions makes it salient that our stipulations of roles need to be conditionalized (like all such stipulations) if they are they are to remain a priori: the most that we can properly stipulate a priori is that 'chance is that quantity, if any, which attaches to competing possibilities and which rational expectations should match', and similarly for the intensity role; it is then an a posteriori matter whether reality supplies anything to fit these stipulations.

concept of an objective chance-like quantity to which rational expectations should conform. This may seem surprising, but doubters will do well to note that, while there was plenty of gambling in antiquity, and certainly enough mathematical sophistication to do the requisite sums, there is no evidence that anybody was able to analyse even the simplest games of chance. ('Someone with only the most modest knowledge of probability mathematics could have won himself the whole of Gaul in a week' [Hacking 1975: 3].)

Let me now turn to the question of role-fillers. Perhaps chance does better than intensity when we consider the actual quantities which fill the respective roles. Maybe the recommendation that we should proportion our expectations to the chancemakers forces itself upon rational thinkers, while the corresponding recommendation about 'intensitymakers' does not.

At first pass, though, this seems unlikely. As so far presented, both orthodoxy and Everett take the relevant role-fillers to be the squared amplitudes of the branches. True, orthodoxy takes these to measure current standings in the competition to become real, whereas Everett takes them to impose a measure over the persistently coexisting branches. But this metaphysical difference was already built into the stipulation of a priori roles, as discussed above. So, if we put this metaphysical difference to one side, it seems that just the same quantity fills the roles on both approaches, which makes it hard to see how rational thought could regard orthodoxy's role-filler as a better guide to expectations than Everett's.

Perhaps Lewis has in mind the specific theory of chancemaking he developed in 'Humean Supervenience Debugged' [1994]. There he argued that chancemakers are patterns of particular fact identified by the Best System. We need not go into details: it is enough to know that the Best System will tend, other things being equal, to equate the chances of outcomes with their omnitemporal relative frequencies, and also to ascribe equal chances to outcomes that display physical symmetries.

This account of chancemaking is motivated by Humean considerations. Lewis wants chances to supervene on matters of particular local fact: two worlds that agree on such facts should not disagree on chances. But his 1994 paper also manifests another motivation. This comes out when he considers the possibility that some special quantity (squared amplitude, perhaps) might play the role of chancemaker. Humeanism by itself does not rule this out: squared amplitudes are as much particular facts as masses are. At this point Lewis appeals to considerations of rationality. He says that he can 'see, dimly but well enough' (p. 484) how frequencies and symmetries can 'constrain rational credence', but cannot 'begin to see' how some primitive theoretical quantity could do this.

Everettians have a strong reply available at this point. They can argue that they are even better placed that the Humean Lewis to explain how their probability-makers 'constrain rational credence'. They can't of course constrain credences by frequencies, since they think that the future will display all frequencies, including rogue frequencies on low intensity branches. But Everettians can relate credences to physical symmetries. Recent work, originating with David Deutsch, argues that, once you accept Everettian metaphysics, then the intensity rule is forced on you by the

physical symmetries of no-collapse quantum theory.[7][7] The argument goes like this: assume only innocuous Everettian variants of standard decision-theoretic constraints on rational preferences, such as transitivity (if you prefer act 1 to act 2, and act 2 to act 3, then prefer act 1 to act 3) and dominance (if act 1's payoffs are greater than act 2's on all branches, then prefer act 1 to act 2); it can then be shown that agents who satisfy these constraints, and act with full quantum knowledge, must always choose as if they are maximizing expected utility with subjective degrees of belief corresponding to the squared amplitudes. In effect, the symmetries of quantum theory plus basic rationality constraints dictate that rational agents will conform to the intensity rule.

There are affinities between the Deutsch approach and 'classical probability theory' which seeks to base probabilities on an assumption of equiprobability for physically symmetrical outcomes, such as heads and tails on a fair coin. But this classical approach runs into trouble within deterministic classical physics, for there have to be symmetry-breaking differences in initial conditions of tossing to explain why coins sometimes come down heads and sometimes tails. So, to reinstate the probabilities for heads and tails within deterministic classical physics, we have to impose a probability distribution over the initial conditions of tossing, and it is by no means obvious that this in turn can be derived solely from physical symmetries.

What about collapse versions of quantum mechanics? These aren't deterministic, so why can't they appeal to the same symmetries as are used by the Everettians? The trouble here is that any serious collapse theory has to add some physical principles to the basic Schrödinger dynamics, in order to explain how and why collapses occur. It would remain to be shown that anything like the Deutsch approach can work given these extra physical principles.

Perhaps collapse chance theorists could take a leaf out of Lewis's Humean book, and somehow use omnitemporal frequencies to show why collapse chances should constrain rational credence. But I must confess I have no clear conception of how orthodox collapse interpretations of quantum mechanics might do this.[8][8]

Let me sum up the analysis of this subsection. We started with Lewis's suspicion that intensity is unlikely to figure 'in any basic principle of rationality', since it is 'a recently discovered and theory-laden magnitude'. On examination, however, there seem no good grounds for this suspicion. For one thing, the a priori concept of chance is itself a relatively recent innovation. Moreover, insofar as basic rationality hinges on frequencies or symmetries, intensity would seem to have a positive advantage, since Deutsch's approach allows Everettians to ground credences in physical symmetries, while it is unclear whether orthodox collapse theories can do anything similar.

More generally, let me now sum up the last three subsections' overall analysis of the relative justifiability of the chance and intensity rules . Lewis initially charged Everettians with positing two mysteries where orthodoxy has only one. But in section

---

[7][7] [Deutsch 1999; Wallace 2002]. See also [Saunders 2003]. I would like to thank David Wallace and Simon Saunders for helpful discussion of the Deutsch approach.
[8][8] Moreover, taking omnitemporal frequencies to be chancemakers generates problems of its own. See [Lewis 1994].

2b I showed that Everettians have no more mysteries than orthodoxy, since they can simply substitute the intensity rule for the chance rule, rather than adding it as an extra. Then in section 2c I argued that it is orthodoxy which has two mysteries, not Everett, since orthodoxy has to square its chance rule with our concern for our unique actual successors, where Everettians recognize no such unique actual successors. Finally, in this subsection I have suggested that Everettians may be left with no mystery at all, since they can justify their intensity rule by appeal to physical symmetries and basic principles of rational preference, while it is unclear whether orthodoxy can do anything similar for its chance rule.

Everettian metaphysics is strange and troubling, where orthodoxy seems familiar and comfortable. But it seems to me that if we consider the matter without prejudice, and make sure we do not mistake familiarity for cogency, it is orthodoxy that is in a tangle with probabilities, not Everett.

3. Life-and-Death Branching

3a. Lewis on Life-and-Death Cases.

In the second half of his paper Lewis focuses on the implications of the no-collapse view for life-or-death situations. These are cases where a conscious creature has some non-zero probability of dying. For example, imagine you are yourself in the position of Schrödinger's cat, and will die if the poison is released.

Lewis argues that in any such case an Everettian should fully expect to survive. After all, the no-collapse metaphysics guarantees that there will be a future branch on which you emerge from the box unscathed, alongside the branch where you die. Moreover, since you will have no experiences on the death branch, it makes no sense for you to expect that branch, so you should have a 100% expectation of surviving.

Lewis generalizes, and concludes that in life-and-death cases we need to adjust the intensity rule, and apportion all our expectation to the branches where we survive. 'When we have life-and-death branching, . . . first discard all the death branches, because there are no minds and no experiences associated with death branches. Only then divide expectations of experience between the remaining branches in proportion to their intensities' (p. 00).

Lewis then considers repeated life-or-death situations. For example, suppose that you are subject to a hundred repetitions of Schrödinger's experiment, being required to re-enter the box each time you survive it. Orthodoxy gives you only a very low chance of coming out of this composite experiment alive. But once more the no-collapse view seems to advise full confidence in survival: after all, at each stage there will be a branch on which you survive, and your cumulative experience of such branches is the only experience you should expect, so you should expect it 100%.

Lewis then makes three further points about repeated life-or-death trials. (1) If Everett is right, we can all expect to live indefinitely. For all causes of death are all probabilistic, and so an Everettian should expect to survive every successive threat of death. (2) If and when you do find that you have repeatedly survived death threats, you will have strong evidence for the no-collapse theory and against orthodoxy. For

orthodoxy implies that such repeated survival is highly unlikely, whereas it is just what you should expect on the no-collapse view. (3) Don't think that the indefinite survival guaranteed by Everett is good news. You may be guaranteed to survive indefinitely, but you should also expect to survive sans friends, sans abilities, sans everything that makes life worth living.

Lewis does not suggest that any of this discredits the no-collapse view. His main concern is not to disprove Everettian metaphysics, but to show it promises a dismal immortality. 'You who bid good riddance to collapse laws . . . should shake in your shoes. Everett's idea is elegant, but heaven forfend it should be true! Sad to say, a reason to wish that it is false is no reason to believe that it is false' (p. 00).

3b. Why Modify the Intensity Rule?

I think Lewis's analysis of life-or-death branching is flawed. There is no reason to modify the intensity rule to deal with life-or-death cases in the first place. And, given this, there is no reason for Everettians to despair about the prospect of indefinitely prolonged misery.

It is true that the no-collapse view implies that we will all have successors who will survive miserably into the indefinite future. But this is no reason for Everettians to feel particularly downcast. For Everettians can also look forward to future branches on which they die at a proper time, along with yet other branches where they die a few years too soon or too late. Taking all these branches together, there is no reason for Everettians to feel any worse about the future than orthodox thinkers: whatever your metaphysics, you should have an infinitesimal expectation that you will survive miserably for an indefinite time, and more reasonable expectations for a normal spectrum of timely and untimely deaths.

Of course, these last remarks assume that Everettians can follow orthodoxy in having normal expectations about futures in which they are dead. And this is precisely what Lewis's 'modified intensity rule' denies. In normal non-life-or-death cases, Everettian expectations quantitatively match those of orthodoxy, despite the divergence in underlying metaphysics. But when it is matter of life-or-death, says Lewis, expectations should come out quite differently, with Everettians distributing the totality of their expectations over those futures where they survive, and discounting altogether futures where they perish.

Still, it is by no means obvious why Everettians should modify their intensity rule in this way. For it seems perfectly open for them to apply the unmodified intensity rule in life-or-death situations, just as elsewhere. If they do this, then they can expect all futures in proportion to their intensities, whether or not those futures contain any of their live successors. For example, even when you know you are about to be the subject in a fifty-fifty Schrödinger's experiment, you should expect a future branch where you perish, to just the same degree as you expect a future branch where you survive.

To bring out the oddity of modifying the intensity rule in the way Lewis suggests, consider the implications for rational action. Rational agents will maximize benefit over all future possibilities weighted by the degree to which they expect those

possibilities. So, if you have a 50% expectation that you will die in Schrödinger's experiment, and regard dying as a very bad thing, then you will have a strong reason not to participate in the experiment. But someone whose expectations are governed by the modified intensity rule will attach a zero expectation to this death branch, and so will not regard the danger of death as any reason to avoid the experiment. More generally, rational agents whose expectations are guided by the modified intensity rule will happily embrace any threat of death, for their only expectation will be of survival.[9]

Why exactly does Lewis think that Everettians should modify their intensity rule in the face of life-or-death situations? It is not clear. He does consider the alternative of keeping the uncorrected intensity rule to govern expectations 'not of experience but of what will happen whether experienced or not'. His response is simply 'But this is not really the intensity rule, which governed only expectations of experience' (footnote 23).

However, this just moves the question around. Why start off in by formulating the intensity rule in terms of 'expectations of experience' in the first place , rather than in terms of 'what will happen whether experienced or not'? When Lewis first introduces the intensity rule, in section 3, it is in the context of theoretical confirmation. Everettians need probabilities so that they can make probabilistic theoretical predictions which might be confirmed or disconfirmed by later experience. As Lewis puts it '. . . we need some new way for no-collapse quantum mechanics to advise us what to <u>expect</u>' (p. 00, my italics). Now, it is true that, as far as theoretical confirmation goes, it is only expectations of <u>experience</u> that matter, for the simple reason that it is only when some such expectation is confirmed or disconfirmed by actual experience that we will get a shift in confirmation. But this is no good reason to formulate the intensity rule in terms of 'expectations of experience', rather than 'expectations of what will happen whether experienced or not'. For 'expectations of experience' are all special cases of 'expectations of what will happen whether experienced or not', and so an intensity rule formulated in the latter terms will still provide everything we need for purposes of theoretical confirmation.

We can put the point like this. Theoretical confirmation is one purpose for which we need expectations. And for this purpose expectations of experience are all we need. But we also need expectations to guide our rational choices, and here an intensity rule formulated solely in terms of expectations of experience will lead us astray. In particular, such an intensity rule will stop us attaching expectations to branches in which we will have no experience, and so will fail to persuade us to avoid dangers of death.[10]

3c There is No Such Thing as a Free Confirmation.

---

[9] Though Lewis himself does not highlight this death-defying consequence of the modified intensity rule, he does refer approvingly to Peter J. Lewis [2000], who focuses on precisely this implication. Note also that David Lewis suggests that, to the extent that you already believe Everett, you will be relatively unworried about risking death in order to further confirm the theory (p. 00).
[10] See [Papineau 2003b] for a critical discussion of some other possible motivations for modifying the intensity rule.

If the intensity rule remains unmodified in life-or-death situations, then Everettians will not be able to confirm their theory in the easy way that Lewis envisages. Lewis's suggestion was that Everettians can confirm their theory simply by repeatedly subjecting themselves to dangers of death and noting that they survive every time. But without the modified intensity rule this doesn't work. Those Everettians who do eventually find that they have survived repeated dangers of death[11] will no longer be able to regard their survival as confirming their no-collapse theory. For a no-collapse theory with an unmodified intensity rule will no longer predict that repeated survival is particularly likely, and so the observation of such repeated survival will add no credence to that theory.

Some readers may be wondering whether Everettians <u>need</u> to modify their intensity rule in order to engineer themselves a free confirmation. Let me conclude my analysis by briefly addressing this issue, as it brings out an interesting aspect of the Everett interpretation. Consider the proposition that, after some repeated life-or-death trial, the future will contain <u>a</u> branch on which I survive. Even with an unmodified intensity rule, this existential proposition is guaranteed to be true by Everett, but very unlikely given orthodoxy. The branch where I survive may have a low intensity compared with the death branches, but Everett still implies with certainty that this branch will exist, where orthodoxy by no means does so. So, when my survivor does find himself in the future, and observes something which Everett says is certain but orthodoxy say is unlikely, it might seem that this by itself will constitute evidence favouring Everett over orthodoxy.

Indeed there is nothing in this reasoning that hinges on <u>death</u> risks, repeated or otherwise. Consider the simple toss of a quantum coin. Everett guarantees that the future will contain a branch with Heads, whereas orthodoxy says this is only 50% probable. So when one of my successors finds that the future does indeed contain Heads, this alone would seem to confirm Everett. (Moreover, this reasoning would seem equally available to my successor who sees Tails.)

There is a flaw in this line of reasoning, however. When you adjust your credence in some theory in the light of new evidence, it is essential that you take into account <u>all</u> your new evidence. Otherwise you can boost the confirmation of the theory simply by focusing on positive evidence and ignoring other evidence. And that is precisely what is going on here. For when I discover that the future contains a survivor of mine, or that it contains a Heads result, this isn't all I discover. In addition I learn that <u>I</u> am the survivor, or that <u>I</u> have observed Heads. This was by no means guaranteed. It was very unlikely that I would survive, rather than end up on a death branch, and it was only 50% likely that I would see Heads. So if we take into account everything that I learn, including these egocentric facts about which branch I am now on, then the no-collapse theory ascribes just the same probability to these facts as orthodoxy, and so receives no greater confirmation when they are observed.

It is true that, from an orthodox point of view, egocentric propositions like <u>I am alive</u>, or <u>I see Heads</u>, are unusual propositions to be the primary bearers of probabilities.[12]

---

[11] Of course, without the modified rule, Everettians will now have good reason to avoid such dangers. However, not all dangers can be avoided.

[12] Orthodoxy does have some room for egocentric probabilities. If I believe that there is an objective chance of p% that I will be transported to Manchester in my sleep, then orthodoxy will advise

But note that probabilities for such egocentric propositions were built into the no-collapse theory from the start. If I am an Everettian, and envisage opening Schrödinger's box, I have no uncertainty about the impersonal structure of the future. I know for sure that it will contain a successor of David Papineau who sees a live cat, and one who sees a dead cat. All I am unsure about is what I will see. Admittedly, Everettians have work to do to articulate the structure of these egocentric probability spaces, and to explain the associated metaphysics of splitting selves.[13][13] Still, given that egocentric probabilities were always at the heart of the no-collapse theory, there seems nothing ad hoc about appealing to these egocentric probabilities in order to get the confirmation relations right.

To sum up, there is no quick route to the confirmation of the no-collapse theory, whichever way we turn the intensity rule. However, Everettians need not feel disappointed that these invitations to free confirmation have been withdrawn. As we have seen, anybody who accepted these invitations would end up committed to fallacious reasoning. In any case, Everttians have no need of confirmational charity. They have support enough in the two considerations that originally motivated the no-collapse theory. The first motivation lies in the theory's direct experimental support: in all circumstances simple enough for the interference effects that demonstrate non-collapse to be detected (like the two-slit experiment) such interference is indeed observed. The second motivation is the inferiority of the theoretical alternatives: all the ad hoc accounts which nevertheless posit collapses (under conditions too complex to test for interference effects) are not only strikingly less elegant than Everett, but also require the rejection of special relativity and the conservation of energy.

As far as empirical confirmation goes, then, the no-collapse theory is as healthy as could be. The experimental evidence leaves no room for anything except highly ad hoc alternatives. If there is reason to resist the no-collapse theory, it can only be that it does not make proper sense, that it somehow undercuts itself by assuming concepts to which it is not entitled. I hope that this paper has at least shown that there is no reason why Everettians are not entitled to the concept of probability.

* * *

Many people have been lessened by David Lewis's early death. The philosophical community in particular has lost one of its guiding lights. Those who attach credence to the no-collapse theory may find some consolation in the thought that there are survivors of the David Lewis that we knew on other branches of reality. These survivors will not all be thriving, as Lewis himself stressed in his lecture. But some of them, perhaps the preponderance, will be in good health and continuing to spread philosophical enlightenment. Still, even for Everettians, such thoughts will provide no more than consolation. If there are other branches of reality, they are permanently

---

me to attach a credence of p% to I am in Manchester when I wake up and have no other information. Note, however, that that this egocentric proposition stands or falls with the non-egocentric proposition David Papineau is in Manchester. The Everett interpretation is different in requiring egocentric probabilities that cannot be so reduced to non-egocentric ones. (Some would argue that even orthodoxy sometimes requires irreducibly egocentric probabilities, as for instance in the 'Sleeping Beauty' puzzle [Elga 2000; Lewis 2001; Arntzenius 2002; Monton 2002]. However, the analysis of this puzzle is by no means clear-cut.)

[13][13] For a start, see [Saunders 1998], especially section 4.3.

inaccessible to us. David Lewis has been lost to our section of history, and we are the worse off for it.

David Papineau
Department of Philosophy
King's College London
WC2R 2LS, UK

David.Papineau@kcl.ac.uk

References

Arntzenius, F. 2002. Reflections on Sleeping Beauty, *Analysis* 62: 53-61.

Deutsch, D. 1999. Quantum Theory of Probability and Decisions, *Proceedings of the Royal Society of London* A455: 3129-37.

Dewitt, B. 1970. Quantum Mechanics and Reality, *Physics Today* 23: 30-40.

Dewitt, B. 1972. The Many-Universes Interpretation of Quantum Mechanics, in *Foundations of Quantum Mechanics* ed. B. d'Espangat, New York: Academic Press.

Elga, A. 2000. Self-Locating Belief and the Sleeping Beauty Problem, *Analysis* 60: 143-7.

Everett, H. 1957. Relative State Formulation of Quantum Mechanics, *Review of Modern Physics* 29: 454-62.

Ghirardi, G., Rimini, A., and Weber, T. 1986. Unified Dynamics for Microscopic and Macroscopic Systems, *Physical Review* D34: 470-91.

Hacking, I. 1965. *The Logic of Statistical Inference*, Cambridge: Cambridge University Press.

Hacking, I. 1975. *The Emergence of Probability*, Cambridge: Cambridge University Press.

Lewis, D, 1980. A Subjectivist Guide to Objective Chance, in *Studies in Inductive Logic and Probability, vol II*, ed. R. Jeffery, Berkeley: University of California Press.

Lewis, D. 1994. Humean Supervenience Debugged, *Mind* 103: 473-90.

Lewis, D. 2001. Sleeping Beauty: Reply to Elga, *Analysis* 61: 171-6.

Lewis, P. 2000. What is it Like to be Schrödinger's Cat?, *Analysis* 60: 22-9.

Monton, B. 2002. Sleeping Beauty and the Forgetful Bayesian. *Analysis* 62: 47-52.

Papineau, D. 1995. Probabilities and the Many Minds Interpretation of Quantum Mechanics, *Analysis* 55: 239-46.

Papineau, D. 1996. Theory-Dependent Terms, *Philosophy of Science* 63: 1-20.

Papineau, D. 2003a (1997). Uncertain Decisions and the Many-Minds Theory of Quantum Mechanics, in *The Roots of Reason* , Oxford: Oxford University Press.

Papineau, D. 2003b. Why You don't Want to Get in the Box with Schrödinger's Cat, *Analysis 63*: 51-7.

Pierce, C. 1923. The Doctrine of Chances, in *Chance Love and Logic,* ed. M. Cohen, New York: Harcourt Brace.

Putnam, H. 1987. *The Many Faces of Realism.* LaSalle, Illinois: Open Court.

Saunders, S. 1998. Time, Quantum Mechanics and Probability. *Synthese* 114: 373-404.

Saunders, S. 2003. Derivation of the Born Rule from Operational Assumptions, *arxiv.org/abs/quant-ph/0211138.*

Wallace, D, 2002. Quantum Probability and Decision Theory, Revisited, *arxiv.org/abs/quant-ph/0211104.*