

A thirder and an Everettian: a reply to Lewis's 'Quantum Sleeping Beauty'

David Papineau & Víctor Durà-Vilà

1. Introduction

Since the publication of Elga's seminal paper in 2000, the Sleeping Beauty paradox has been the source of much discussion, particularly in this journal. Over the past few decades the Everettian interpretation of quantum mechanics¹ has also been much debated. There is an interesting connection between the way these two topics raise issues about *subjective probability assignments*.

This connection is often alluded to, but as far as we know Peter J. Lewis's 'Quantum Sleeping Beauty' (2007a) is the first attempt to examine it explicitly. Lewis claims that the two debates are not independent: to be specific, he argues that accepting the Everettian interpretation of quantum mechanics requires you to be a 'halfer' about Sleeping Beauty, in opposition to the more widely accepted 'thirder' solution.

This paper will argue that Lewis is wrong. Everettians do not have to be halfers. It is perfectly cogent to be both an Everettian and a thirder.

2. Elga's original puzzle

The Sleeping Beauty puzzle was introduced by Elga (2000: 143) as follows:

Some researchers are going to put you to sleep [on Sunday]. During the two days that your sleep will last, they will briefly wake you up either once [on Monday] or twice [on Monday and Tuesday], depending on the toss of a fair coin (Heads: once; Tails: twice). After each waking, they will put you to back to sleep with a drug that makes you forget that waking.

When you are first awakened, to what degree ought you believe that the outcome of the coin toss is Heads?

The apparent paradox arises because there are seemingly good reasons to answer both $1/2$ and $1/3$ to the question posed by Elga. On the one hand, given that the coin is fair, you clearly ought to attach a credence of $1/2$ to Heads before being put to sleep, and since you do not acquire any new information on being wakened, there seems no reason not to stick to this. On the other hand, if the experiment were repeated many times, $2/3$ of the awakenings would follow Tails and only $1/3$ would follow Heads, and this seems to argue for attaching a credence of $1/3$ to Heads (Elga 2000: 143-44).

A survey of the different arguments for halving and thirding is beyond the scope of this paper. We need only note that the overwhelming majority of philosophers are thirders (including *inter alia* Elga (2000), Vaidman and Saunders (2001), Dorr (2002), Monton (2002), Arntzenius (2003), Hitchcock (2004), Horgan (2004), (2007), Weintraub (2004) and Dieks (2007)). As far as we know, only David Lewis (2001) is a definite halfer in print. Bradley (2003), Bradley and Leitgeb (2006), White (2006), and Pust (2008) raise challenges of varying importance to thirders.

¹ See Wallace 2002 and references therein for the contemporary version of Everett's (1957) interpretation of quantum mechanics.

3. The Everettian interpretation of quantum mechanics and subjective probability

Suppose that there is a spin-1/2 particle in an eigenstate of z-spin. According to the Everettian interpretation of quantum mechanics, if the x-spin is measured, then fission or branching takes place.² The measurement results in two branches of reality, with the result spin-up in one branch and spin-down in the other. Moreover, if there is an observer of this result, then this observer branches too, with one successor seeing spin-up and the other spin-down.

What degree of belief ought the observer to attach to seeing spin-up before the split? This question raises a number of issues for Everettians (see §7), but they all agree that in this symmetrical situation the pre-split observer should adopt a credence of 1/2 for spin-up (and likewise for spin-down).

Now consider this case, which will be central in what follows. Everything is as above, except that the result of the spin measurement is temporarily hidden after the branching. In this case too all Everettians agree that the post-split successors should assign a 1/2 credence for spin-up until the result is revealed.

4. Simplified Sleeping Beauty and the Everettian analogy

Lewis (2007a) begins with a *simplified* Sleeping Beauty case. This does away with the coin toss. Beauty is simply told that she will be woken up on Monday, that her memory of this will be erased, and that she will then be woken up again on Tuesday. Therefore, on Sunday Beauty ‘should assign a probability of 1 to both “I will wake up on Monday” and “I will wake up on Tuesday”’ (Lewis 2007a: 62). Moreover, upon her first awakening she should assign a probability of 1/2 to ‘Today is Monday’ (and likewise for Tuesday).

Lewis observes that there are some striking similarities between this simplified Sleeping Beauty case and the Everettian set-up where the result of the spin measurement is concealed from the post-split observer. First, in both cases the two later selves are psychologically continuous with the initial self. Second, the two later selves are psychologically discontinuous from each other: any new memories formed by the one are not shared by the other. Finally, in both cases ‘the agent gets lost in [the] branching structure’ (Lewis 2007a: 60): in the simplified Sleeping Beauty case, you will not know which day it is, and in the Everettian case, you will not know which branch you are on until the result of the measurement is revealed.

We can illustrate these analogies diagrammatically. Figures 1 and 2 show the most natural schematization of the two scenarios, while, Figure 3 shows a ‘self-centred’ or ‘phenomenological’ view of Beauty when she wakes up unaware of whether it is Monday or Tuesday.

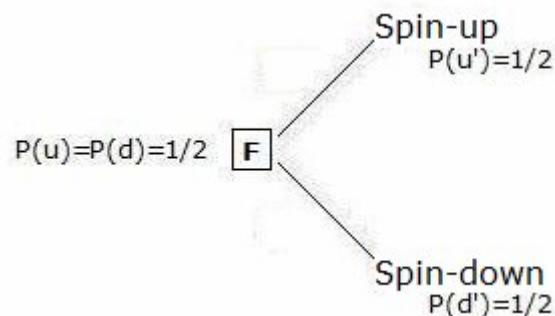


Figure 1: Everettian fission. **F**: fission, u: ‘The electron will be spin-up’, d: ‘The electron will be spin-down’, u’: ‘The electron here is spin-up’, d’: ‘The electron here is spin-down’.

² We are using here the same example as Lewis (2007a: 59-60).

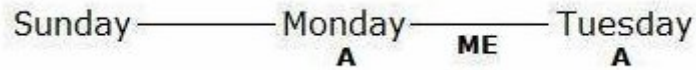


Figure 2: simplified Sleeping Beauty case. A: awakening, ME: memory-erasure

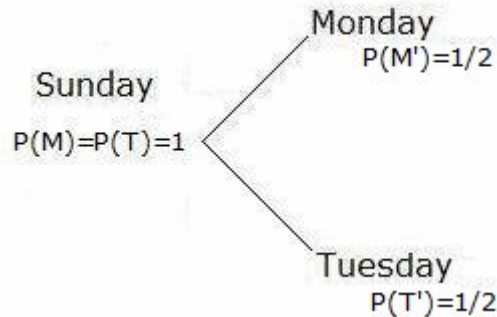


Figure 3: simplified Sleeping Beauty case, phenomenological view. M: 'I will be woken on Monday', T: 'I will be woken on Tuesday', M': 'It is Monday here', T': 'It is Tuesday here'.

5. Lewis's argument

Lewis's argument hinges on the thought that, since the simplified Sleeping Beauty and the Everettian set-up are so structurally similar, the same rational degrees of belief ought to apply in both cases.

As schematized above, the two cases differ in the prior degrees of belief ascribed to the relevant outcomes. Where Beauty ascribes a credence of 1 to both 'I will wake up on Monday' and 'I will wake up on Tuesday' before she is put to sleep on Sunday, the Everettian observer ascribes 1/2 to 'I will see spin-up' and 'I will see spin-down' before splitting takes place.

So Lewis infers that either Beauty ought to have prior degrees of belief of 1/2, or the pre-split Everettian observer ought to have prior degrees of belief of 1. Since he takes it to be essential to the Everettian programme that the pre-split Everettian observer will have prior degrees of belief of 1/2, he concludes that Everettians are forced to hold that Sleeping Beauty ought to attach a prior credence of 1/2 to both 'I will wake up on Monday' and 'I will wake up on Tuesday'.

Given this move, Lewis's conclusion that Everettians have to be halfers quickly follows. The full Sleeping Beauty puzzle can be viewed as an initial coin toss, followed either by the simplified Sleeping Beauty case (if Tails on the initial toss) or by an awakening on Monday alone (if Heads). But Lewis has in effect already concluded that Everettians must treat the simplified Sleeping Beauty case just as if she were going to be awakened on *either* Monday *or* Tuesday depending on the outcome of a fair coin toss. So he infers that Everettians must treat the full Sleeping Beauty puzzle just as if it involved *two* successive coin tosses, as in Figure 4.

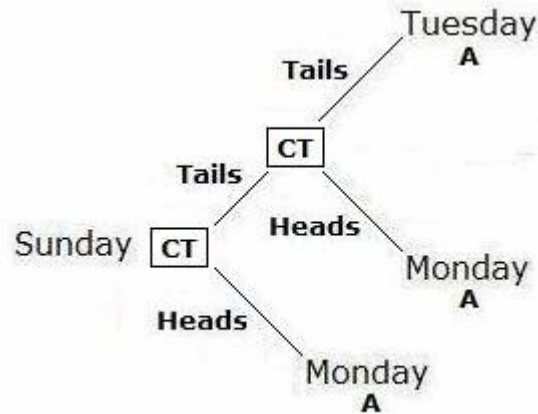


Figure 4: two coin toss Sleeping Beauty case. CT: coin toss, A: awakening

Now, it is uncontested that in a double-toss set-up like this, you ought to be a halfer about the initial coin toss. You are going to be awakened just once whether that initial toss is Heads or Tails. So there is no line of argument which suggests, as in the original full Sleeping Beauty paradox, that subjects will be awakened twice as often when the initial coin toss was Tails, and so should take their being awakened as some sort of evidence for Tails. Being awakened after a double coin toss as in Figure 4 cannot in any way be construed as carrying any information about the initial toss, and so must leave the credence in Tails at 1/2.

6. Criticism of Lewis's argument

Lewis's argument is only as good as his crucial claim that Everettians must treat the simplified Sleeping Beauty case and the Everettian spin measurement analogously. But he offers no good reason for this. True, there are the structural analogies he points to (see §4). But against that can be put the basic point that for Everettians the metaphysics of the simplified Sleeping Beauty set-up is quite different from that of the spin measurement.

There is only one branch of reality at issue in the simplified Sleeping Beauty case, in which she is awoken on *both* Monday *and* Tuesday. By contrast, Everettians hold that there are two branches of reality after the spin measurement, each with only one result. Since Everettians are committed to non-unitary credences only in respect of futures that genuinely branch, they can restrict initial credences of 1/2 to the spin-measuring set-up. In the simplified Sleeping Beauty case there is only one future branch, in which both the Monday and Tuesday awakening will occur for sure, so they should continue to attach a credence of one to these events.

In short, the analogous phenomenology of self-location does not override the fundamental metaphysical differences between the spin-measurement and the simplified Sleeping Beauty set-up. The self-centred perceptions may be similar, but the metaphysical differences still dictate a different assignment of Everettian credences.

If Lewis is wrong about what Everettians should say about the simplified Sleeping Beauty case, then he is also wrong to claim that they should treat the full Sleeping Beauty set-up as a double coin toss. Everettians should simply insist that there is no real chanciness on the Tails branch. Once the coin turns out to be Tails, Beauty will be woken up *both* on Monday *and* on Tuesday. Everettians have no reason to accept Lewis's claim that for them this is equivalent to her being 'woken up *either* on Monday *or* on Tuesday' with a 1/2 chance of each (Lewis 2007a: 63, our emphasis). So his argument that Everettians must be halfers falls away. They can view the full Sleeping Beauty paradox just like everybody else, and so have just as much reason to be thirders.

7. Vaidman on quantum credences

There is one possible line of argument that might bolster Lewis's case. There is no indication in his paper that he would endorse this line of argument, but it will be worth discussing briefly.

Let us go back to the question of Everettian credences about future branches. We said earlier that Everettians all agree that in the spin measurement case the right credences are 1/2 for spin-up and 1/2 for spin-down. But of course there is an obvious worry about their ascription of such non-unitary credences to their putative branching futures. Since they know that the future is certainly going to contain all the branches, why should they expect any of them to a degree other than one?

Different Everettians offer different answers to this challenge. Some hold that non-unitary Everettian credences require some kind of ignorance about the future (Saunders (1998), Tappenden (2000), Wallace (2006) and Vaidman (1998), (2002a)). Others hold that they are consistent with full knowledge of the future (Papineau (2004) and Greaves (2004)). Of interest here is a specific version of the former strategy due to Lev Vaidnam.³

Vaidman's approach hinges on the kind of case where the result of a quantum measurement is temporarily concealed from the post-branching observers. These observers are genuinely ignorant about which branch they are on. Accordingly, their non-unitary credences can be regarded as a normal reflection of this ignorance. Vaidman then argues that since there is no substantial epistemic difference between these observers and their pre-branching ancestor, the ancestor ought to attach the same non-unitary credences to the relevant branches.

This line of thought substantially strengthens the analogy between the spin-measurement case and the simplified Sleeping Beauty case. If Vaidman is right, the pre-split non-unitary credences in the spin-measurement case are a reflection of the way post-split observers can be ignorant about their self-location. If this is really the rationale for the Everettian non-unitary credences, then perhaps it does indeed carry over to simplified Sleeping Beauty case. The awakened Sleeping Beauty is ignorant about her self-location, yet she is not substantially epistemic different from her Sunday self. Vaidman's line of argument thus suggests that her Sunday self ought to have non-unitary initial credences too.⁴

We are not sure whether or not Vaidman's rationale really does carry over from spin-measurements to Sleeping Beauty cases. But suppose for the sake of the argument that it does. Given this, Lewis can then appeal to Vaidman to bolster his claim that Everettians ought to be halfers.

Still, the natural response is that this is a problem for Vaidman's specific approach to non-unitary Everettian credences, rather than a problem for Everettianism per se. It cannot be right to say that Beauty should have non-unitary credences about her future wakings in the simplified Sleeping Beauty case. If Vaidman's account implies this, then so much the worse for Vaidman's account.

After all, as observed above, Vaidman's approach is just one among a number of Everettian options. None of the other options support Lewis's analogy between Everettian quantum measurements and Sleeping Beauty. If Vaidman's approach does indeed validate the analogy, then this is simply a reason for preferring these other options to Vaidman's.⁵

*King's College London
Strand, London WC2R 2LS, UK
david.papineau@kcl.ac.uk*

victor.dura_vila@kcl.ac.uk

³ For discussions of Vaidman's proposal, see Wallace 2006: 678-80 and Lewis 2007b: 10-11.

⁴ It must be noted, however, that Vaidman is a thirder (Vaidman and Saunders 2001; Vaidman 2002b).

⁵ This paper has benefited from discussions with Peter J. Lewis, Darren Bradley and Paul Tappenden.

References

- Arntzenius, F. 2003. Some problems for conditionalization and reflection. *The Journal of Philosophy* 100: 356–70.
- Bradley, D. 2003. Sleeping Beauty: a note on Dorr's argument for 1/3. *Analysis* 63: 266–68.
- Bradley, D. and H. Leitgeb. 2006. When betting odds and credences come apart: more worries for Dutch book arguments. *Analysis* 66: 119–27.
- Dieks, D. 2007. Reasoning about the future: Doom and Beauty. *Synthese* 156: 427–39.
- Dorr, C. 2002. Sleeping Beauty: in defence of Elga. *Analysis* 62: 292–96.
- Elga, A. 2000. Self-locating belief and the Sleeping Beauty problem. *Analysis* 60: 143–47.
- Everett, H. 1957. 'Relative state' formulation of quantum mechanics. *Reviews of Modern Physics* 29: 454–62.
- Greaves, H. 2004. Understanding Deutsch's probability in a deterministic multiverse. *Studies in History and Philosophy of Modern Physics* 35: 423–56.
- Hitchcock, C. 2004. Beauty and the bets. *Synthese* 139: 405–20.
- Horgan, T. 2004. Sleeping Beauty awakened: new odds at the dawn of the new day. *Analysis* 64: 10–21.
- Horgan, T. 2007. Synchronic Bayesian updating and the generalized Sleeping Beauty problem. *Analysis* 67: 50–59.
- Lewis, D. 2001. Sleeping Beauty: reply to Elga. *Analysis* 61: 171–76.
- Lewis, P. J. 2007a. Quantum Sleeping Beauty. *Analysis* 67: 59–65.
- Lewis, P. J. 2007b. Uncertainty and probability for branching selves. *Studies in the History and Philosophy of Modern Physics* 38: 1–14.
- Monton, B. 2002. Sleeping Beauty and the forgetful Bayesian. *Analysis* 62: 47–53.
- Papineau, D. 2004. David Lewis and Schrödinger's cat. *Australasian Journal of Philosophy* 82: 153–69.
- Saunders, S. 1998. Time, quantum mechanics and probability. *Synthese* 114: 373–404.
- Tappenden, P. 2000. Identity and probability in Everett's multiverse. *British Journal for the Philosophy of Science* 51: 99–114.
- Pust, J. 2008. Horgan on Sleeping Beauty. *Synthese* 160: 97–101.
- Vaidman, L. 1998. On schizophrenic experiences of the neutron or why we should believe in the many-worlds interpretation of quantum theory. *International Studies in the Philosophy of Science* 12: 245–61.
- Vaidman, L. and S. Saunders. 2001. On Sleeping Beauty controversy. <http://philsci-archive.pitt.edu/archive/00000324/>. Accessed 17 December 2007.
- Vaidman, L. 2002a. Many-worlds interpretation of quantum mechanics. In *The Stanford Encyclopedia of Philosophy*, ed. E. N. Zalta. <http://plato.stanford.edu/entries/qm-manyworlds/>. Accessed 17 December 2007.
- Vaidman, L. 2002b. Probability and the many worlds interpretation of quantum mechanics. In *Quantum Theory: Reconsideration of Foundations*, ed. A. Khrennikov, 407–22. Växjö, Sweden: Växjö University Press.
- Wallace, D. 2002. Worlds in the Everett interpretation. *Studies in History and Philosophy of Modern Physics* 33: 637–61.
- Wallace, D. 2006. Epistemology quantized: circumstances in which we should come to believe in the Everett interpretation. *British Journal for the Philosophy of Science* 57: 655–89.
- Weintraub, R. 2004. Sleeping Beauty: a simple solution. *Analysis* 64: 8–10.
- White, R. 2006. The generalized Sleeping Beauty problem: a challenge for thirders. *Analysis* 66: 114–19.