# Sleeping Beauty: a note on Dorr's argument for 1/3

DARREN BRADLEY

Beauty is about to be drugged, rendering her unconscious for a long time. During that time she will be awakened briefly, either once (on Monday) or twice (on Monday and Tuesday). The number of awakenings depends on the toss of a fair coin: if the result is Tails, she is awakened twice: if Heads, once. The nature of the drug is that she will not remember being awake. In particular, when she is awakened, she will not know whether it is Monday or Tuesday. Upon awakening on Monday, what should her degree of belief be that the coin landed Heads?

The paradox is that there are compelling arguments for the answers of both 1/2 and 1/3. Cian Dorr gives an argument by analogy that the correct

answer is 1/3. I will argue that his analogy is dissimilar in a crucial respect and as such is unhelpful to resolving the problem.

Dorr presents the following variation on the story(2002: 293–94):

> Beauty is about to go to sleep for a long time. During that time she will be awakened briefly, twice, on Monday and Tuesday. A fair coin will be tossed, and what Beauty can remember will depend on the result. If the coin lands Tails, she will be given the same amnesia-inducing drug as was used in the original experiment. She will wake up twice, unable to tell if it is the first or second awakening. But if the coin lands Heads, she will be given a weaker amnesia-inducing drug, which merely *delays* the onset of memories from the previous day, rather than destroying them completely. If Beauty receives this weaker drug, the first minute of her awakening on Tuesday will be just as it would have been if she had received the stronger drug, but after that the memories of Monday's awakening will come flooding back. She will realise that it is Tuesday and that the outcome of the toss must have been Heads.

Let $P_-$ be Beauty's credence function immediately after being woken on Monday in the variant case. We must assign credence for the following 4 hypotheses:

H1   The coin lands Heads and it's Monday
H2   The coin lands Heads and it's Tuesday
T1   The coin lands Tails and it's Monday
T2   The coin lands Tails and it's Tuesday

How should this credence be distributed? Dorr claims, and I agree, that the only plausible answer is: $P_-(H1) = P_-(T1) = P_-(H2) = P_-(T2) = 1/4$.

Let $P$ be Beauty's credence function after a minute has passed on Monday. Assume she has not had the flooding back of memories she would have experienced had H2 been true, so $P(H2) = 0$. Nothing in her experience during the first minute does anything to discriminate between the other three hypotheses. So the ratio of her credence in H1, T1 and T2 will remain unchanged. Hence $P(H1) = P(T1) = P(T2) = 1/3$; so $P(Heads) = P(H1) = 1/3$. I agree with this reasoning.

Dorr then argues that there are no relevant differences between his variant case and the original case with respect to $P(H1)$. I disagree. There is a crucial difference between Beauty after a minute in the variant case and Beauty after a minute in the original case. In the variant case, a certain possibility has been eliminated. It *could* have turned out that it was Heads and Tuesday. Not only could it have been Heads and Tuesday, Beauty could have *known* it was Heads and Tuesday. She would have known this for certain if she had suddenly got a flood of memories from the previous day.

This flood of memories would have made Heads more likely. The absence of this flood makes Heads less likely. That is, she can conditionalize on the absence of a flood of memories and conclude that Heads is less likely.

While she was waiting for a possible flood of memories in that first minute, she had no new evidence about how the coin landed. If she got a flood of memories, she would increase the probability that the coin landed Heads (to 1). So in the absence of a flood of memories, she would have to decrease the probability that the coin fell Heads.

In the original case, there is no such possibility. There is no possibility that Beauty will find out for certain that it fell Heads. So she cannot update on the lack of such information. Unlike in the variant case, Beauty has no non-indexical information to conditionalize on.

Dorr asks why this delay of a minute, the only difference between the original Sleeping Beauty and the variant case should make any difference to subjective probabilities. I answer that in the variant case a hypothesis that could have been eliminated has failed to be (H2). There is no equivalent of this is in the original case. The variant case is uninformative.[1]

*Stanford University*
*Stanford, CA 94305, USA*
*dbradley@stanford.edu*

*References*

Dorr, C. 2002. Sleeping Beauty: in Defence of Elga. *Analysis* 62: 292–96.

---

[1] I would like to thank Branden Fitelson for helpful discussion.