An Embarrassment for Double-Halfers

Michael G. Titelbaum
University of Wisconsin–Madison

“Double-halfers” think that throughout the Sleeping Beauty Problem, Beauty should keep her credence that a fair coin flip came up heads equal to 1/2. I introduce a new wrinkle to the problem that shows even double-halfers can’t keep Beauty’s credences equal to the objective chances for all coin-flip propositions. This leaves no way to deny that self-locating information generates an unexpected kind of inadmissible evidence.

Keywords chance; credence; Principal Principle; self-location; Sleeping Beauty Problem

DOI:10.1002/tht.21

The Sleeping Beauty Problem: Beauty volunteers for an epistemology experiment. She will be put to sleep Sunday night; the experimenters will flip a fair coin. If the coin comes up heads, they will awaken her Monday morning, chat with her for a bit, tell her that it’s Monday, then put her back to sleep. If the coin comes up tails, they will engage in the same Monday process then erase all her memories of the Monday awakening, awaken her Tuesday morning, chat with her for a bit, tell her it’s Tuesday, then put her back to sleep.

All this is explained to Beauty, then she goes to sleep. When she awakens Monday morning, how confident should she be that the coin came up heads?

After introducing this problem to the philosophical literature, Adam Elga (2000) noted that when Beauty awakens Monday morning, she entertains three possibilities:

- it’s Monday and the flip outcome is heads
- it’s Monday and the flip outcome is tails
- it’s Tuesday and the flip outcome is tails

(Beauty doesn’t awaken Tuesday if the coin comes up heads.) Elga argues that Beauty should distribute her credence equally among these three possibilities. Since the coin comes up heads on only the first possibility, Elga is a “thirder” about the Sleeping Beauty Problem—he thinks Beauty should be 1/3 confident when she awakens Monday morning that the coin came up heads.

Correspondence to: E-mail: titelbaum@gmail.com
There’s something odd about this credence assignment. When Beauty awakens Monday morning she is certain that the coin flip was fair, and she hasn’t been informed of its outcome. So it feels like all this business about waking up, going to sleep, and what day it is doesn’t make it rational for her to deviate her confidence in heads from 1/2. This intuition drives many philosophers to take a “halfer” position on the Sleeping Beauty Problem, demanding that Beauty assign 1/2 credence to heads upon awakening. For that reason I’ll call it the Original Halfer Intuition. 

Elga pointed out that it’s difficult for the halfer to maintain this intuition consistently. To see why, note that the outcome of the coin flip doesn’t change anything the experimenters do until after Beauty goes to sleep on Monday—their experimental protocol is the same on either outcome up to that point. So we can move the crucial coin flip to Monday night. Since this doesn’t change anything that happens to Beauty, the move shouldn’t change anything about Beauty’s required Monday morning confidence in heads. But now consider Beauty’s credence in heads once she’s been told it’s Monday, a time we’ll call Monday afternoon. At that point Beauty is certain that the coin is fair, is certain that it hasn’t been flipped yet, and seems to have no other information about the flip’s outcome. So a devotee of the Original Intuition should require Beauty to assign a 1/2 credence to heads. To dramatize the point, we can imagine the experimenters putting the coin in Beauty’s hand and saying, “This is the coin we’re going to flip in ten minutes. It’s fair—it has a 1/2 objective chance of coming up heads. And however the flip comes out, its outcome has no influence on your present condition.” At that point there certainly seems to be no good reason for Beauty to deviate her credence in heads from its objective chance.

Yet keeping credences and chances together on Monday afternoon is difficult for the halfer. Due to some mathematics I’ll leave for a footnote, as long as we make the reasonable assumption that Beauty assigns a nonzero credence when she first awakens to its being Tuesday, the halfer takes Beauty to assign a credence greater than 1/2 to the coin’s coming up heads conditional on its being Monday. This means that if Beauty updates her credences between Monday morning and Monday afternoon by conditionalizing on the fact that it’s Monday, her Monday-afternoon unconditional credence in heads will be greater than 1/2. That assignment is in tension with the halfers’ Original Intuition.

David Lewis (2001) understood this implication and tried gamely to defend the resulting Monday-afternoon halfer credences. Lewis’s preferred direct inference principle relating credences and chances, the Principal Principle (Lewis 1980), states (roughly) that an agent must match her credences to what she is certain are the objective chances unless she has “inadmissible evidence”. While Lewis didn’t think Beauty has any inadmissible evidence on Monday morning that would permit her to deviate her credence in heads at that point from the objective chances, he tried to find some inadmissible evidence Beauty possesses on Monday afternoon that would make his version of halfing (with a greater-than-1/2 heads credence on Monday afternoon) work. But even Lewis admitted that doing so would require a “novel and surprising” move (2001, p. 175).

But many halfers are unwilling to make this novel move with him, largely because it violates their Original Intuition—that all this sleeping/waking/day-of-the-week stuff
doesn’t make it okay for Beauty to assign credence other than 1/2 to the outcome of an objectively fair coin flip. Halfers adhere to this intuition by defending a “double-halfer” position, on which Beauty is required to be 1/2 confident in heads both before and after she learns that it’s Monday. Double-halfers pull this off by denying that Beauty’s Monday-afternoon unconditional credence in heads should equal her Monday-morning credence in heads conditional on its being Monday. In other words, they deny that when Beauty learns it’s Monday she should update her credences by conditionalizing on that fact.6

By sacrificing this bit of conditionalization, double-halfers keep Beauty’s credence in a fair coin flip equal to 1/2 all Monday long. But I will now show that their victory is short-lived. Suppose the experimenters enjoy flipping the coin so much that they decide to flip it once more on Tuesday night. The coin remains fair, the Tuesday flip has no impact on anything having to do with Beauty (it’s just an idle coin flip), and the Tuesday flip will be performed whether Beauty awakens that day or not. Assume also that Beauty is informed on Sunday night that the extra Tuesday flip will occur. When Beauty awakens Monday morning (uncertain what day it is), how confident should she be in the proposition that today’s coin flip comes up heads?

By the probability calculus, this credence must be greater than Beauty’s credence that the Monday coin flip comes up heads. To see why, recall that only one of Elga’s three possibilities had the Monday flip coming up heads: what we’ll now call “Monday-and-Monday-heads” to indicate that on this eventuality today is Monday and the Monday flip comes out heads. So Beauty’s credence that the Monday flip comes up heads is just her credence in Monday-and-Monday-heads. Next notice that today’s coin flip comes up heads just in case it’s Monday and the Monday flip comes up heads or it’s Tuesday and the Tuesday flip comes up heads. So Beauty’s credence that today’s flip comes up heads is the sum of her credences in Monday-and-Monday-heads and Tuesday-and-Tuesday-heads. But Beauty’s credence in Monday-and-Monday-heads just is her credence that the Monday flip comes up heads, so as long as she has nonzero credence in Tuesday-and-Tuesday-heads (which she should7), her credence that today’s coin flip comes up heads must be greater than her credence that Monday’s flip does.

The halfer wants Beauty upon first awakening to assign a 1/2 credence that the Monday coin flip comes up heads. But then Beauty’s credence that today’s flip comes up heads must be greater than 1/2—and that’s embarrassing. The double-halfer’s guiding intuition is that all this sleeping/waking/day-of-the-week stuff doesn’t permit Beauty to assign an unconditional credence other than 1/2 to the outcome of a fair coin flip. Beauty is certain that today’s coin is fair, is certain that it hasn’t been flipped yet, and has no information about its outcome. Yet the halfer needs her to be greater than 1/2 confident that today’s flip will come up heads.8

Notice that this embarrassment stems entirely from Beauty’s credences when she first awakens Monday morning; no conditionalization step was needed. Simply by taking the halfer’s preferred Monday-morning distribution and applying the probability calculus, we’ve shown that the halfer needs Beauty to be more than 1/2 confident that today’s flip will come up heads. A number of authors have argued that in cases involving self-locating
(or “centered”) propositions standard updating by conditionalization fails to generate rational constraints on credences. But no one has argued that self-location causes the probability calculus to fail.9

The double-halfer might still suggest that there’s something strange about credences in self-locating propositions like the proposition that today’s coin flip comes up heads, and this strangeness makes Beauty’s (admittedly counterintuitive) greater-than-1/2 credence in that proposition acceptable. But this is ridiculous. Imagine the experimenters put the coin in Beauty’s hand and say, “This is the coin we’re going to flip in ten minutes. It’s fair—it has a 1/2 objective chance of coming up heads. And however the flip comes out, its outcome has no influence on your present condition. Heck, if you like you can be the one to flip it.” Standing there with the coin in her hand, Beauty is supposed to be more than fifty percent confident that it’ll come up heads?10

Despite double-halfers’ best attempts, there is no way to keep credences equal to objective chances throughout the Sleeping Beauty Problem. At some point (whether it be on Monday morning or on Monday afternoon), with respect to some proposition (whether it be that Monday’s flip comes up heads or that today’s flip does), Beauty’s self-locating information must provide her with a novel and surprising kind of inadmissible evidence.11

Notes

1 Of course if Beauty is awakened Tuesday, when she is told it’s Tuesday that will send her credence in heads to 0. But telling Beauty it’s Tuesday is a way of informing her about the outcome of the coin flip.

2 Notice that the Original Halfer Intuition is an intuition about synchronic relations—it holds that Beauty’s self-locating information on Monday morning doesn’t authorize her to deviate her credence in heads at that time away from what’s she certain at that time is the objective chance of heads. Halfers sometimes appeal to a different, diachronic intuition that when Beauty awakens on Monday morning and learns the self-locating information that it’s now Monday or Tuesday, that shouldn’t change her credence in heads from its value on Sunday night—which everyone in the debate agrees should be 1/2. In Titelbaum (2008), I characterized such diachronic intuitions about the relevance of self-locating information to non-self-locating propositions (such as propositions about the outcome of the coin flip) as instances of a “Relevance-Limiting Thesis”. That thesis has subsequently been disproven by examples having nothing to do with the Sleeping Beauty Problem—see Bradley (2011, Section 9), Titelbaum (Forthcoming-a, Chapter 10), and Titelbaum (Forthcoming-b, Section 4). Thus the diachronic motivation for halving seems to me to have been decisively undermined, and the Original Halfer Intuition is the most popular remaining motivation for halving.

3 Assigning a nonzero credence to its being Tuesday means assigning a nonzero credence to Tuesday-and-tails. If Monday-and-heads receives 1/2 credence the tails possibilities sum to 1/2, so Monday-and-tails receives credence less than 1/2. By the standard Ratio Formula for conditional credences, Beauty’s credence in heads conditional on Monday is her credence in Monday-and-heads divided by her credence in Monday. Since the latter is the sum of her credences in Monday-and-heads and Monday-and-tails, the only way to make the ratio 1/2
or less would be to make Beauty’s credence in Monday-and-tails at least 1/2. But we just said
that Beauty’s credence in Monday-and-tails is less than 1/2.

4 Thirders have no trouble achieving a Monday-afternoon credence in heads of 1/2. Going
back to Elga’s equal distribution over Beauty’s possibilities, you can calculate that Beauty’s
Monday-morning credence in heads conditional on its being Monday is exactly 1/2.

5 For attempts to make sense of Lewis’s surprising move, see Bradley (2011) and Titelbaum
(Forthcoming-a, Chapter 9).

6 There are two ways to make this work: either one denies that conditionalization applies to all
updates involving self-locating credences (and in particular to the Sleeping Beauty update in
question)—see Meacham (2008), Meacham (2010), Briggs (2010), Cozic (2011), and Pust
(Forthcoming)—or one argues that Beauty should generate Monday-afternoon credences by
conditionalizing Monday-morning credences on something other than the fact that it’s
Monday—see Bostrom (2007).

7 Anyone who wants credences to mirror chances will admit that as long as Beauty has some
positive credence that it’s Tuesday, she should have some positive credence that it’s Tuesday
and the Tuesday flip comes up heads. So the only challenge to a positive
Tuesday-and-Tuesday-heads value is from authors like Hawley (Forthcoming) and Schwarz
(2012) who suggest that when Beauty awakens Monday morning she should be certain that
it’s Monday. But this suggestion is awfully implausible—it requires Beauty to have zero
confidence in a possibility (that it’s Tuesday) that hasn’t been ruled out by her evidence.
(Among other things, this suggestion grossly violates the Regularity Principle.) Moreover, if
Beauty is certain on Monday morning that it’s Monday, the distinction between thirding and
double-halfing disappears; both thirders and double-halfers agree that Beauty should be 1/2
certain in the Monday coin’s coming up heads once she’s certain that it’s Monday. But
double-halfers make a big deal of their disagreement with thirders!

8 The thirder doesn’t have this problem. Like the halfer, the thirder has Beauty assign a greater
credence upon awakening to today’s flip coming up heads than she does to Monday’s flip
coming up heads. But for the thirder the latter value is 1/3, so there’s no problem with the
former value’s being 1/2.

9 All the double-halfers listed in note 6 are perfectly happy applying the probability axioms to
self-locating credences. The only author I’m aware of who might object at this stage would be
Peter Lewis (2010). Lewis maintains a double-halfer stance by denying my earlier claim that
Beauty’s Monday-morning credence that the Monday flip comes up heads is just her credence
in Monday-and-Monday-heads. Lewis thinks the former credence should be 1/2 and the latter
1/3. He defends this position by arguing that no legitimate Dutch Book can be constructed
against an agent who assigns those values. But Dutch Book or no Dutch Book, I have a hard
time believing it could be rational for Beauty to be more confident of one possibility than
another when she is certain that the former will eventuate just in case the latter does.

10 The double-halfers’ preferred Monday-morning credence distribution for Beauty, combined
with the probability calculus, generates a credence that today’s flip will come up heads which
directly conflicts with the Original Intuition. Now not all double-halfers are motivated by
what I’ve called the Original Halfer Intuition. But all of them (assuming they apply the
probability calculus to subjective credences) are committed to the same ridiculous
consequence that Beauty’s credence that today’s coin flip will come up heads should be
greater than 1/2.
My Tuesday-flip version of the Sleeping Beauty Problem developed out of discussions with Wolfgang Schwarz about his work on decision theory and the Absentminded Driver Paradox (a forerunner of Sleeping Beauty). I am grateful to Schwarz, Rachael Briggs, Chris Meacham, and Joel Pust for discussion of the example.

References


Hawley, P. (Forthcoming). Inertia, Optimism and Beauty. *Noûs*.


