Sleeping Beauty on Monty Hall

Michel Janssen* and Sergio Pernice†

December 21, 2019

Abstract

Inspired by the Monty Hall Problem and a popular simple solution to it, we present a simple solution to the notorious Sleeping Beauty Problem. We replace the awakenings of Sleeping Beauty by contestants on a game show like Monty Hall's and we increase the number of awakenings/contestants in the same way that the number of doors in the Monty Hall Problem is increased to make it easier to see what the solution to the problem is. We show that the Sleeping Beauty Problem and variations on it can be solved through simple applications of Bayes's theorem. This means that we will phrase our analysis in terms of credences or degrees of belief. We will also rephrase our analysis, however, in terms of relative frequencies. Overall, our paper is intended to showcase, in a simple yet non-trivial example, the efficacy of a tried-and-true strategy for addressing problems in philosophy of science, i.e., develop a model for the problem and vary its parameters. Given that the Sleeping Beauty Problem, much more so than the Monty Hall Problem, challenges the intuitions about probabilities of many when they first encounter it, the application of this strategy to this conundrum, we believe, is pedagogically useful.

Keywords: Sleeping Beauty Problem; Monty Hall Problem; probability; Bayesian; frequentist.

*Program in the History of Science, Technology, and Medicine, University of Minnesota; email: janss011@umn.edu

†Universidad del Cema, Buenos Aires; email: sp@cema.edu.ar

1 Warm-up exercise: The Three Stooges on Monty Hall

Consider the following puzzle. In a special edition of his famous game show, "Let's Make a Deal", Monty Hall calls the Three Stooges to the stage and has them collectively pick one of three doors, D_1, D_2 or D_3. Behind one are two checks for a thousand dollars each, behind the other two is a goat. The Three Stooges pick D_2. Before the show goes to a commercial break, Monty tells his special guests that either one or two of them will be called back after the break. If they were wrong, only one of them will return (but he doesn't tell them which one!); if they were right, the other two will. During the break, the Three Stooges are made to take a nap backstage. When the show resumes, they are sound asleep.

If the checks are *not* behind D_2, Monty wakes up Curly and brings him back to the stage, making sure he has no idea whether he is the first or the second one to be woken up and brought back. As usual, Monty opens either D_1 or D_3 (whichever one has a goat behind it) and offers Curly to switch from D_2 to the other door that remains unopened. If the door he chooses has the two checks behind it, he gets one of them. If not, he goes home empty-handed.

If the checks *are* behind D_2, Monty goes through this same routine twice, with Moe and Larry (not necessarily in that order), the only other difference being that he now has a choice whether to open D_1 or D_3. Monty makes sure that neither of them finds out whether they were woken up and brought back first or second (at least not until Monty opens the door with the checks behind it, one of which may be gone at that point).

What is a Stooge to do in this predicament? Should he switch? Should he stay with the door they initially picked? Should he be indifferent between staying and switching?[1]

2 Introduction: Monty Hall and Sleeping Beauty

The puzzle in Section 1 combines elements of two well-known puzzles challenging our intuitions about probability: the Monty Hall Problem and the Sleeping Beauty

[1] Three honors students at the University of Minnesota (Rose Adams, Jamie DeBruyckere, and Terrance Gray) worked with one of us (MJ) to make a video posing this riddle. We used drawings by Laurent Taudin as the basis for the animation. Scott Spicer and Charlie Heinz of Library Media Services at the University of Minnesota helped the students record the voice-overs and the soundtrack. The result was posted on YouTube: See https://www.youtube.com/watch?v=5U5-6CWB0vQ.

Problem.[2] The solution of the Monty Hall Problem is no longer controversial. A contestant in a normal episode of "Let's Make a Deal" should always take Monty Hall up on his offer to switch. Since Monty Hall never opens the door with the prize behind it and thus has to know which door that is, we need to assume that he offers contestants to switch regardless of which door they initially picked but that assumption is routinely granted.

Given this assumption, the intuition that the opening of a door and the offer to switch are just for dramatic effect and do not affect the contestant's chances of winning is simply wrong. The opening of one of the doors provides the contestant with important information. A simple and (judging by its ubiquity on the web) effective way to make the point is the following (though the key assumption mentioned above is not always spelled out). Initially, the contestant only has a 1/3 chance of picking the right door and a 2/3 chance of picking the wrong one. Suppose he (or she) switches. If he was right the first time, he will now be wrong. If he was wrong the first time, he will now be right. So he now has a 2/3 chance of being right and only a 1/3 chance of being wrong. In a slogan, he flips the odds by switching. A device often used to shore up one's intuitions in this case is to increase the number of doors. If there are n doors ($n \geq 3$), the contestant initially has a $1/n$ chance of picking the right one and a $(n-1)/n$ chance of picking the wrong one. After the contestant has made her (or his) initial choice, Monty Hall opens all but one of the remaining doors and offers her to switch. Once again, the contestant flips the initial odds by switching. By switching, she effectively guesses that the prize is not behind the door she initially picked but behind any one of the other $n-1$ doors, all but one of which Monty Hall has meanwhile opened for her.

Unlike the Monty Hall Problem, the Sleeping Beauty Problem remains controversial. The problem is essentially the following. Sleeping Beauty is told that a fair coin will be tossed after she's been put to sleep and that, when she is woken up, she will be asked what her degree of belief is that the coin came up heads. It depends on the outcome of that coin toss, however, how many times she is asked. If the coin comes up heads, she will only be woken up and asked once. If the coin comes up tails, she will be woken up and asked twice. The first time she's woken up after the coin comes up tails she is given some amnesia drug so that she won't remember the second time that she's been woken up before and asked the same question. Every precaution is taken to make sure that Sleeping Beauty, when she wakes up, cannot tell whether her current awakening is the one after the coin came up heads or one of the two after the coin came up tails. What should Sleeping Beauty's degree of belief

[2] There is a vast literature on both. Good places to start are the wikipedia entries for these two problems.

be upon being awakened that the coin came up heads?

One knee-jerk response is that, no matter how often Sleeping Beauty is put to sleep, woken up and drugged, the probability that a fair coin comes up heads is and remains 1/2. Therefore, her answer every time she is asked should be 1/2. Another knee-jerk response is that, if the coin comes up tails, Sleeping Beauty is twice as likely to be asked than if the coin comes up heads. Therefore, her answer every time she is asked should be 1/3. Those who think the answer is 1/2 are known as halfers. Those who think the answer is 1/3 are known as thirders. The debate between halfers and thirders has long moved beyond this clash of intuitions upon first encountering the problem. We strongly suspect that it persists mainly because it is ambiguous exactly what Sleeping Beauty is being asked.[3]

To get a better handle on the Sleeping Beauty Problem and inspired by the Monty Hall Problem, we design a game-show proxy for it, in which potential contestants take over the role of potential awakenings. Our analysis will show that both halfers and thirders are right, depending on how one interprets the question Sleeping Beauty is asked. That said, our game-show proxy will be much more congenial to thirders than to halfers as it implements the interpretation of the question they would consider to be the interesting one. *Should the information that Sleeping Beauty is given ahead of time about what will happen depending on the outcome of one toss of a fair coin change her degree of belief upon being awakened that this particular coin toss resulted in heads?* Upon a little reflection, both halfers and thirders will agree that if *that* is the question, Sleeping Beauty would be just as mistaken to ignore this information as contestants on "Let's Make a Deal" would be to ignore the information Monty Hall is giving them by opening one of the three doors.

Just as Sleeping Beauty knows how a coin toss will decide how many of her potential awakenings will become actual awakenings, potential contestants in our game-show proxy know how a coin toss will decide how many of them will become actual contestants. Both Sleeping Beauty and our potential contestants are asked to

[3]The Sleeping Beauty Problem is due to Elga (2000), who took the thirder position. His paper provoked a response by Lewis (2001), who took the halfer position. For an up-to-date review of the professional literature on the Sleeping Beauty Problem, see Cisewski *et al.* (2016). As Albert Einstein wrote to the Italian mathematician Tullio Levi-Civita on August 2, 1917: "It must be a pleasure to ride through these fields on the steed of real mathematics, while the likes of us have to muddle through on foot" (*Es muss hübsch sein, auf dem Gaul der eigentlichen Mathematik durch diese Gefilde zu reiten, während unsereiner sich zu Fuss durchhelfen muss*) (Schulmann *et al.*, 1998, Doc. 368). Cisewski 2016 *et al.* throw everything and the kitchen sink at Sleeping Beauty. The only mathematical tool we will use is Bayes's theorem. For the use of Bayesian reasoning in connection with both the Sleeping Beauty Problem and the Monty Hall Problem, see the dissertation by Bradley (2007).

assess (one way or another) the probability that this coin toss resulted in heads when woken up or selected as a contestant, respectively. Using Bayes's rule, they should update their degree of belief that the coin came up heads in light of the information about how potential awakenings/contestants become actual awakenings/contestants. In the spirit of a time-honored Bayesian tradition (based on the principle that you put your money where your mouth is), we will cash out the degrees of belief of our game-show contestants in terms of betting behavior.[4] If our contestants update according to Bayes's rule (and the same is true for Sleeping Beauty), their prior degrees of belief, determined by the symmetry of the coin, get replaced by posterior degrees of belief that are determined by the numbers of contestants (awakenings) selected for the two possible outcomes of the coin toss—1 for heads, 2 for tails in the original problem but those numbers can be chosen arbitrarily. Although the paper is mostly written in terms of credences or degrees of belief, the analysis does not require this interpretation of probabilities. It can easily be rephrased in terms of relative frequencies in long series of repetitions of our gameshow.

Similar game-show proxies can be designed for variations on the Sleeping Beauty Problem in which the coin toss is replaced by a different stochastic experiment with an arbitrary number of outcomes and arbitrary numbers of candidates (awakenings) for different outcomes. As in the case of the Monty Hall Problem, a good way to shore up one's intuitions in dealing with the Sleeping Beauty Problem, as we will see in Sections 3–4, is to vary these parameters. In Section 5, we illustrate this strategy for getting a handle on the Sleeping Beauty Problem by designing a game-show proxy for a variation on the problem that combines it with the Monty Hall Problem (for arbitrary numbers of doors and awakenings). The scenario with the Three Stooges in Section 1 is essentially our game-show proxy for this combined problem for three doors and three awakenings. In Section 6, we use the game-show proxy for the combined problem to tweak the so-called 50-50 lifeline in the TV show "Who wants to be a millionaire?" In this game-show, unlike the preceding ones, the degrees of belief used as priors in our Bayesian analysis no longer reflect symmetries of physical systems. The analysis, however, can still easily be rephrased in frequentist terms.

[4] In true Bayesian fashion, Hitchcock (2004) offers a diachronic Dutch Book argument in support of the thirder position.

3 Game-show proxies for (variations on) the Sleeping Beauty Problem

We introduce a new game show loosely modeled on Monty Hall's "Let's make a deal". The show starts with the host selecting (not necessarily randomly) a pool of N candidates or potential contestants from the audience. These candidates are carefully sequestered before the host proceeds to perform a simple stochastic experiment with n equiprobable outcomes O_i ($i = 1, \ldots, n$): flipping coins, rolling fair dice, spinning a roulette wheel, drawing numbered balls from a bingo cage, drawing cards from a deck, etc. Different stochastic devices can be used in different versions or different episodes of the show. Depending on the outcome of the experiment, the host then calls up, one by one, one or more of the N candidates to become actual contestants on the show. Each contestant is given a certain amount of money, say a thousand dollars, to place a bet, at even odds,[5] on one of the n possible outcomes of the experiment the host just performed (that very experiment, not a repetition of it). If they are right, they go home with two thousand dollars; if they are wrong, they go home empty-handed.

If contestants were randomly selected from the initial pool of N candidates, their chances of winning the bet would be $1/n$ regardless of which outcome O_i they choose to put their money on. This probability is determined simply by symmetries of the physical object(s) used in the stochastic experiment. As long as he or she does not know anything about the selection procedure, a contestant's degree of belief that the outcome was O_i should be $1/n$ for all i. In Bayesian terms, these should be his or her prior degrees of belief or *priors* for short:

$$\Pr(O_i) = \frac{1}{n}. \tag{1}$$

In our game show, however, the selection procedure is anything but random and all candidates know that it is not. Here is how it works. At the beginning of the show the candidates are divided into groups assigned to different outcomes. The sizes of these groups will be different for different outcomes. The candidates are told the sizes of the groups for all outcomes but they are not told in which group they are. Candidates become contestants if and only if they happen to be in the group for the actual outcome of the stochastic experiment. The host will call up all members of this group, one by one, to place their bet.

[5] The restriction to bets at even odds can be relaxed but this would unnecessarily complicate the analysis.

Figure 1: Game show with Monty Hall, eight potential contestants, eight soundproof booths and a lottery ball machine as the stochastic device (Drawing by Laurent Taudin).

To make sure that a candidate, when selected as a contestant, knows neither the outcome of the stochastic experiment nor whether any other candidates have already been called up as contestants, all candidates are put into individual soundproof booths after the host has explained to them how contestants will be selected but before he carries out the stochastic experiment (see Fig. 1). When a booth's door is closed, the candidate inside cannot detect in any way what is going on outside of it. The candidates are given the sizes of the groups for all outcomes before they are sent to their booths.

Each booth is equipped with a buzzer to let the candidate inside know that she has been selected as a contestant. The candidate-turned-contestant then leaves her booth, closes the door behind her, places the bet and finds out whether she lost or won. Unless she is the final contestant, she is only allowed to watch the rest of the show from a place where the contestants coming after her cannot see or hear her (or detect her presence in any other way). Once all members of the group of candidates selected as contestants have placed their bets, the remaining candidates are allowed to leave their booths and the show is over.

How can the candidates take advantage of the information they are given about the selection of contestants to improve their chances of winning their bets when selected as a contestant?

Let N_i be the number of candidates in the group assigned to the outcome O_i. These N_i's can be any non-negative integers that add up to the total number of candidates N. Of course, the number of candidates will itself be a number chosen by the producers of the show. Note that one or more N_i's are allowed to be zero. If $N_j = 0$ and O_j happens to be the result of the stochastic experiment, no candidates will become contestants. That is not a problem. The host can simply repeat the stochastic experiment until he gets an outcome O_k for which $N_k \neq 0$. We can think of the instances in which the stochastic experiment produces a result O_j for which $N_j = 0$ as runs of the game where the game is over before contestants are called to the stage.

The candidates can use Bayes's rule to update the priors in Eq. (1) to take into account the information they are given about how candidates are selected to become contestants and use their posterior degrees of belief (or *posteriors* for short) when deciding which outcome to put their money on if and when actually selected as a contestant. They can do this updating as soon as they have been given the group sizes N_i, even though their beliefs do not really matter unless and until they become contestants (see Fig. 2).

We will first do this Bayesian updating for the simple case in which the stochastic experiment is a single coin toss. We then generalize the results to arbitrary stochastic experiments with n equiprobable outcomes.

For a single toss of a fair coin, we have

$$n = 2, \quad O_1 = H, \quad O_2 = T, \quad N_1 = N_H, \quad N_2 = N_T, \quad N = N_H + N_T. \tag{2}$$

H and T stand for 'Heads' and 'Tails', respectively. N_H and N_T are the number of candidates in the groups for heads and tails, respectively. These should add up to the total number of candidates, N. The prior degrees of belief for any candidate that the coin will come up heads or tails, respectively, are:

$$\Pr(H) = \Pr(T) = \frac{1}{2}. \tag{3}$$

We need to calculate the posteriors $\Pr(H \,|\, A)$ and $\Pr(T \,|\, A)$, where A stands for

Figure 2: A potential contestant in our game show sitting in her booth, contemplating what to do when chosen as an actual contestant (Drawing by Laurent Taudin).

actual as opposed to *potential* contestant and is defined as

$$A \;=\; \text{I, a candidate, have been selected as a contestant, aware of the protocol for selecting candidates to become contestants.}$$

Since we have equal priors, we can use the simple version of Bayes's rule to calculate the posteriors:

$$\Pr(H \,|\, A) = \frac{\Pr(A \,|\, H)}{\Pr(A \,|\, H) + \Pr(A \,|\, T)}, \quad \Pr(T \,|\, A) = \frac{\Pr(A \,|\, T)}{\Pr(A \,|\, H) + \Pr(A \,|\, T)}. \qquad (4)$$

The evaluation of the posteriors thus comes down to the evaluation of the *likelihoods* $\Pr(A \,|\, H)$ and $\Pr(A \,|\, T)$.

One might be tempted to set $\Pr(A \,|\, H) = \Pr(A \,|\, T) = 1/2$ on the argument that the probability of any candidate being selected as a contestant is $1/2$ regardless of the outcome of the coin toss. In that case, we would also have $\Pr(H \,|\, A) = \Pr(T \,|\, A) = 1/2$. We would be ignoring the information, however, that different numbers of candidates will become contestants depending on the outcome of the coin toss, which is part of the statement A. That information thus also needs to be taken into account.

If the coin came up heads (H) and the candidate is selected as a contestant (which is entailed by A), she must be one of the members of the group for heads. The likelihood $\Pr(A \mid H)$ is thus equal to the probability that she belongs to that group. Great precautions have been taken to make sure that, as far as she can tell, she could be any one of the N candidates. The probability that she (or any other candidate) belongs to the group for heads is therefore simply the number of candidates N_H in that group divided by the total number of candidates N. A similar argument can be given for $\Pr(A \mid T)$. Hence

$$\Pr(A \mid H) = \frac{N_H}{N}, \qquad \Pr(A \mid T) = \frac{N_T}{N}. \tag{5}$$

Inserting these expressions in Eq. (4), we arrive at

$$\Pr(H \mid A) = \frac{N_H}{N_H + N_T} = \frac{N_H}{N}, \quad \Pr(T \mid A) = \frac{N_T}{N_H + N_T} = \frac{N_T}{N}. \tag{6}$$

These posteriors replace the equal priors in Eq. (3). In the case of equal priors (we will see in Section 4 that we have to be more careful in the case of unequal priors), the procedure for selecting contestants from an initial pool of candidates thus has the effect of replacing the priors, determined by the symmetry of the coin, by posteriors that are determined by the choice of the sizes N_H and N_T of the two groups into which the initial pool of N candidates are divided. What Eq. (6) shows then is that the producers of the show can basically change the (equal) priors to any posteriors they like.

Since the posteriors in Eq. (6) only depend on N_H and N_T, candidates will already know as they are sitting in their booths waiting for the buzzer to go off, whether they will choose heads or tails, should they be called up as contestants. If $N_H > N_T$, they choose heads. If $N_T > N_H$, they choose tails. If $N_H = N_T = N/2$, they get no information that would increase their chances of winning their bet. For $N_H = N$ (or, similarly, $N_T = N$), finally, the game becomes trivial: all contestants would be guaranteed to win. If the coin comes up heads, the host calls up all contestants; if it comes up tails, he just flips it again.

The same conclusions can be reached when we conceive of probabilities in terms of relative frequencies rather than degrees of belief. Imagine that this version of our game show (with the same coin, the same pool of N potential contestants, and the same numbers N_H and N_T) is repeated X times (with X a very large integer). Now put yourself in the shoes of an arbitrary potential contestant. You expect to be selected as a contestant in about N_H/N of the roughly $X/2$ runs of the show in which the coin will come up heads and in about N_T/N of the roughly $X/2$ runs in which the coin will come up tails (if $N_H = 0$ or $N_T = 0$ the show will be over without any contestants being selected in roughly $X/2$ runs and you will be selected as a contestant in the roughly $X/2$ remaining runs). Overall, you thus expect to be selected as a contestant in about

$$\frac{N_H}{N}\frac{X}{2} + \frac{N_T}{N}\frac{X}{2} \tag{7}$$

runs. Since $N_H + N_T = N$, Eq. (7) confirms that, overall, you expect to be selected as a contestant in about half of the X runs. However, Eq. (7) also shows that, if $N_T > N_H$, you expect to be a candidate in runs where the coin comes up tails more often than in runs in which the coin comes up heads. In other words, that you (or any other potential contestant) should use Eq. (6) and set the probabilities that a fair coin came up heads or tails equal to $N_H/N < \frac{1}{2}$ and $N_T/N > \frac{1}{2}$, respectively, is simply to take into account this *sampling bias*.

These results can readily be extended to arbitrary stochastic devices with n equiprobable outcomes O_i ($i = 1, \ldots, n$). The prior degrees of belief for any candidate that the experiment will result in O_i are

$$\Pr(O_i) = \frac{1}{n}. \tag{8}$$

For each outcome O_i there will be $N_i \geq 0$ potential candidates (the host will repeat the stochastic experiment until he gets an outcome O_j for which $N_j \neq 0$). Since we have equal priors, we can use the simple version of Bayes's rule to calculate the posteriors:

$$\Pr(O_i \,|\, A) = \frac{\Pr(A \,|\, O_i)}{\sum_{j=1}^{n} \Pr(A \,|\, O_j)}. \tag{9}$$

The candidate's knowledge about the selection procedure tells him or her that the likelihoods are given by (cf. Eq. (5) and the reasoning leading up to it)

$$\Pr(A \,|\, O_i) = \frac{N_i}{N}. \tag{10}$$

Inserting this expression for the likelihoods on the right-hand of Eq. (9), we see that Eq. (6) generalizes to:

$$\Pr(O_i \,|\, A) = \frac{N_i}{N}. \tag{11}$$

Once again, we see that the protocol for selecting contestants from an initial pool of candidates has the effect of replacing (equal) priors (see Eq. (8)), determined by symmetries of the stochastic device used, by posteriors (see Eq. (11)) that are determined by the choice of the numbers N_i, which are subject only to the requirement that they add up to N, the number of candidates we started out with.

Like Eq. (6), Eq. (11) can be seen as expressing a certain sampling bias. Imagine that this version of our game show (with the same stochastic device, the same pool of N potential contestants, and the same set of numbers $\{N_i\}$) is repeated X times (with X a very large integer). Put yourself in the shoes of an arbitrary potential contestant in these X runs of the game. Suppose there are $m < n$ outcomes O_j for which $N_j = 0$. That means that in roughly $m(X/n)$ runs of the game, no contestants are selected. In the remaining $(n-m)(X/n)$ runs, you will be selected as a contestant in about N_i/N of the roughly X/n runs in which the outcome of the stochastic experiment is O_i. Overall, you will thus be selected in about

$$\sum_{i=1}^{n} \frac{N_i}{N}\frac{X}{n} \tag{12}$$

runs. Since $\sum_{i=1}^{n} N_i = N$, Eq. (12) confirms that you expect to be selected as a contestant in about X/n runs overall. However, Eq. (12) also shows that, if $N_k > N_l > 0$, you expect to be a candidate in runs with outcome O_k more often than in runs with outcome O_l. In other words, that you (or any other potential contestant) should use Eq. (11) and choose a value other than $1/n$ for the probability that the stochastic experiment under consideration results in one of its n equiprobable outcomes is simply to take into account this sampling bias.

4 Game-show proxy for the original version of the Sleeping Beauty Problem

Our proxy for the Sleeping Beauty Problem is a special case of the game show analyzed with malice aforethought in Eqs. (2)–(7) in Section 3. It is a version or episode of the show in which the stochastic experiment is a single toss of a fair coin. We analyzed this game show for an arbitrary number of candidates N divided into groups for heads and tails, containing N_H and N_T candidates, respectively. To turn this into a proxy for the Sleeping Beauty Problem, we set $N = 3$ (the minimum number of candidates for which this version of the show is non-trivial) and $N_H = 1$.

The parameters N_H, N_T and N introduced in Eq. (2) thus have the values:

$$N_H = 1, \quad N_T = 2, \quad N = 3. \tag{13}$$

Candidates (potential contestants) are the analogues of potential awakenings. The one candidate selected to become an actual contestant if the coin comes up heads is the analogue of the one potential awakening of Sleeping Beauty that becomes an actual awakening if the coin comes up heads. The two contestants selected after the coin comes up tails are the analogues of the two potential awakenings that become actual awakenings if the coin comes up tails. The drugging of Sleeping Beauty ensures that during subsequent awakenings she knows as little about preceding awakenings as any contestant coming out of his or her booth knows about other contestants in the game-show proxy for the Sleeping Beauty Problem.

It may seem that there is still an important difference between potential contestants and potential awakenings. It is easy to tell potential contestants apart: we could use their Social Security Numbers (SSNs), for instance. Every potential contestant will presumably know at all times what his or her SSN is. How can we tell Sleeping Beauty's potential awakenings apart? How can Sleeping Beauty herself tell her own potential awakenings apart? Fortunately, these questions have easy answers. We can specify, for any potential awakening, at what time and on what date it will happen if it will happen at all (allowing ten minutes or so for each awakening). We then put a clock that displays both time and date in the room where Sleeping Beauty is awakened so that she can tell, every time she is woken up, which potential awakening has just become an actual awakening. Just as we had no need to refer to a potential candidate's SSN, we will have no need to refer to date and time of a potential awakening. The point is that, at least in principle, individuating potential awakenings is no more problematic than individuating potential contestants.[6]

This way of providing potential awakenings with a time stamp should not be conflated with the time stamps used in the standard version of the Sleeping Beauty Problem. There Sleeping Beauty is told ahead of time that she will be woken up a second time on Tuesday if the coin comes up tails. In that case, Sleeping Beauty, when awakened, can obviously not be allowed to find out what day of the week it is (at least not until she has answered the question what her degree of belief is that the coin came up heads). In our version, she is told only that two of the three potential awakenings (labeled by the time and date they will happen) will become actual awakenings if the coin comes up tails. She is not told which two. In that case, the date and the time she is awakened are as irrelevant to Sleeping Beauty's assessment of her degree of belief that the coin came up heads as the SSN of a potential contestant in our game-show proxy is to his or her calculation of the posteriors $\Pr(H \mid A)$ and $\Pr(T \mid A)$ (see Eqs. (4)–(6)).

[6] Our substitution of 'contestants' for 'awakenings' likewise suggests that the notion of "self-locating beliefs" invoked by Elga (2000) is irrelevant to the solution of the Sleeping Beauty Problem.

Sleeping Beauty has no reason to doubt that the coin used to determine whether she will be woken up once or twice is fair. In other words, her prior degree of belief that the coin will come up heads should be

$$\Pr(H) = \frac{1}{2}. \tag{14}$$

What should Sleeping Beauty's degree of belief be, once she is woken up, that the particular toss of this coin that decided how many times she would be awakened resulted in heads? In Bayesian terms, what is her posterior $\Pr(H \,|\, A)$? Here A (which, as before, stands for *actual* as opposed to *potential*) is defined as

A = I, Sleeping Beauty, have been awakened, aware of the protocol governing which of my potential awakenings become actual awakenings.

With this reinterpretation of A and using the values of the parameters in Eq. (13), we can use the first half of Eq. (6) for Sleeping Beauty's posterior degree of belief that the coin came up heads:

$$\Pr(H \,|\, A) = \frac{N_H}{N} = \frac{1}{3}. \tag{15}$$

This follows directly from Eq. (5) for the likelihoods $\Pr(A \,|\, H)$ and $\Pr(A \,|\, T)$.

Since this is the crux of the matter, we transfer the reasoning leading up to Eq. (5) back to the Sleeping Beauty Problem. We do so for arbitrary values of N_H and N. If the coin came up heads (H) and she has been awakened aware of the protocol for selecting awakenings (A), Sleeping Beauty knows that her present awakening must be one of the N_H potential awakenings that become actual awakenings if the coin came up heads. The likelihood $\Pr(A \,|\, H)$ is thus equal to Sleeping Beauty's degree of belief that her current awakening is one of those N_H potential awakenings. Great precautions have been taken to make sure that Sleeping Beauty, when she awakes, has no way of telling whether her current awakening is one of the N_H potential awakenings should the coin come up heads or one of the N_T potential awakenings should the coin come up tails. The probability that her current awakening is one of the former is therefore simply the number N_H of such potential awakenings divided by the total number N of all potential awakenings.

In Section 2, we noted that the Monty Hall Problem becomes more intuitive if the number of doors is increased. The solution to the Sleeping Beauty Problem likewise becomes more intuitive if the number of potential awakenings or, in our game-proxy for the problem, the number of potential contestants is increased. Instead of the values in Eq. (13), we choose

$$N_H = 1, \quad N_T = 99, \quad N = 100. \tag{16}$$

In this case, Eq. (15) tells us that Sleeping Beauty's degree of belief that the coin came up heads plummets from its initial value of 1/2 to just 1/100.

Like Eq. (6) and Eq. (11), Eq. (15) can be seen as expressing a certain sampling bias. Imagine we repeat the experiment with Sleeping Beauty X times where X is some very large integer. In the roughly $X/2$ runs in which the coin comes up heads, the total number of awakenings of Sleeping Beauty will be $N_H(X/2)$. In the roughly $X/2$ runs in which the coin will come up tails, the total number of awakenings will be $N_T(X/2)$. In all runs combined, the total number of awakenings will thus be about $N_H(X/2) + N_T(X/2) = N(X/2)$. If $N_T > N_H$, more of these awakenings occur during runs in which the coin comes up tails than during runs in which the coin comes up heads. That Sleeping Beauty uses Eq. (15) to set the probability that a fair coin came up heads equal to $N_H/N \neq \frac{1}{2}$ is simply to take into account this sampling bias.

Eq. (15) shows that, *as long as our game show can be used as a proxy for the Sleeping Beauty Problem*, the thirders are right and the halfers are wrong. There is a simple way, however, of reconciling the two positions. Once again imagine that the experiment with Sleeping Beauty is repeated many times. Ask Sleeping Beauty to answer the following pair of questions:[7]

Question #1: What is the probability that in an arbitrarily chosen run the coin comes up heads? **Answer**: 1/2.

Question #2: What is the probability that an arbitrarily chosen awakening happens during a run in which the coin comes up heads? **Answer**: N_H/N (which works out to 1/3 for the values given in Eq. (13)).

One can debate—and commentators have (see notes 2 and 3)—whether question #1 or question #2 is the more natural way of interpreting the question in the original Sleeping Beauty Problem and whether the problem doesn't become trivial once you accept that halfers and thirders are just interpreting it differently. We will not get into those debates. We simply note that, although both questions are relevant

[7]This way of reconciling the positions of halfers and thirders follows Groisman (2008):

> [A]n automatic device tosses a fair coin, if the coin lands 'Tails' the device puts two red balls in a box, if the coin lands 'Heads' it puts one green ball in the box. The device repeats this procedure a large number of times ... The core of the whole confusion is that we tend to regard [the probability that] 'A (one) green ball is put in the box' and [the probability that] 'A green ball is picked out from the box' as equivalent (Groisman, 2008, p. 4).

to our game-show proxy (the answer to question #2 presupposes the answer to question #1), it is the second question that makes this game show and others like it intrinsically interesting or, at least, amusing to contemplate.

5 Switch or Stay? Sleeping Beauty on Monty Hall

Suppose Sleeping Beauty is a contestant in a special episode of Monty Hall's "Let's Make a Deal". Monty asks her choose between doors D_1, D_2 and D_3. Behind one, he tells her, are two checks for a thousand dollars each; behind the other two is a goat. Sleeping Beauty chooses D_2. Monty offers her the usual deal but with a twist. The twist is that he will only open one of the other doors (as usual we assume that Monty knows which door has the checks behind it and that he will always open a door with a goat behind it) and give her the opportunity to switch *after* she's been put under and woken up, possibly twice (in which case he will have to administer an amnesia drug the first time). Every time she is awakened she is offered the same deal and Monty Hall will open one of the doors (not necessarily the same one each time) and offer her to switch. If Sleeping Beauty was right the first time and the checks are behind D_2, she will be woken up twice. If she was wrong the first time and the checks are behind D_1 or D_3, she will only be woken up once. That she might have to be drugged once, Monty tells her, is well worth it, as she will get one of the checks each time she picks the right door and could thus conceivably walk away with two thousand dollars.

Sleeping Beauty takes the deal. When awakened, should she switch doors, should she stay with the one she originally picked, or should she be indifferent between switching and staying?

The special episode of "Let's Make a Deal" with the Three Stooges in Section 1 is essentially our game-show proxy for this Sleeping Beauty on Monty Hall Problem. The only difference is that instead of being made to take a nap backstage like the Three Stooges, our contestants will all be put in booths like the ones we used in the game show introduced in Section 3. Fig. 3 illustrates this particular format of the new game show.

As in Section 1, we assume the three candidates initially pick D_2. They are now divided into two groups. Two of them are assigned to D_2, one of them is assigned to the other two doors. This makes the format of the game show significantly different from the one we introduced and analyzed in Section 3. There all candidates were assigned to one of n *equiprobable* outcomes O_i. In this case, they are assigned to *non*-equiprobable outcomes. This means that we can no longer use the simple form of Bayes's rule (for equal priors). We need to use the general form (for unequal priors).

Figure 3: Monty Hall opens one of the doors with a goat behind it for one of the contestants in an episode of "Let's Make a Deal" with a twist inspired by the Sleeping Beauty Problem (Drawing by Laurent Taudin).

Other than that, however, the analysis is completely analogous to the analysis in Section 3.

We have two (non-equiprobable) outcomes,

$$D_2 = \text{Checks behind door \#2}, \quad \text{not-}D_2 = \text{Checks not behind door \#2}. \tag{17}$$

The priors are:

$$\Pr(D_2) = \frac{1}{3}, \quad \Pr(\text{not-}D_2) = \frac{2}{3}. \tag{18}$$

We have three candidates divided into two groups (one of two, one of one) assigned to these two outcomes:

$$N_{D_2} = 2, \quad N_{\text{not-}D_2} = 1, \quad N = N_{D_2} + N_{\text{not-}D_2} = 3. \tag{19}$$

The likelihoods are (cf. Eq. (5) and the reasoning leading up to it):

$$\Pr(A \,|\, D_2) = \frac{N_{D_2}}{N} = \frac{2}{3}, \quad \Pr(A \,|\, \text{not-}D_2) = \frac{N_{\text{not-}D_2}}{N} = \frac{1}{3}, \tag{20}$$

where A is defined as in Section 3 as "I, a candidate, have been selected as a contestant, aware of the protocol for selecting candidates to becomes contestants." The posterior $\Pr(D_2 \,|\, A)$ is given by (the general form of) Bayes's rule:

$$\Pr(D_2 \,|\, A) = \frac{\Pr(A \,|\, D_2)\,\Pr(D_2)}{\Pr(A \,|\, D_2)\,\Pr(D_2) + \Pr(A \,|\, \text{not-}D_2)\,\Pr(\text{not-}D_2)}. \tag{21}$$

The posterior $\Pr(\text{not-}D_2 \,|\, A)$ has the same denominator but $\Pr(A \,|\, \text{not-}D_2)\,\Pr(\text{not-}D_2)$ in the numerator. From Eqs. (18) and (20), we see that

$$\Pr(A \,|\, D_2)\,\Pr(D_2) = \Pr(A \,|\, \text{not-}D_2)\,\Pr(\text{not-}D_2) = \frac{2}{9}, \tag{22}$$

which means that

$$\Pr(D_2 \,|\, A) = \Pr(\text{not-}D_2 \,|\, A) = \frac{1}{2}. \tag{23}$$

The sampling bias introduced by our peculiar procedure of selecting contestants from an initial pool of candidates thus wipes out the advantage of switching doors a contestant would have had in a regular episode of "Let's Make a Deal". Replacing the three potential contestants by three potential awakenings, we conclude that Sleeping Beauty should be indifferent between staying with the door she initially picked and switching to one other door that remains unopened.

Of course, we can just as easily create a version of the Sleeping Beauty on Monty Hall Problem in which she increases her chances of picking the right door from 1/3 to 2/3 by staying. To do that, we replace the values of N_{D_2}, $N_{\text{not-}D_2}$ and N in Eq. (19) by

$$N_{D_2} = 4, \quad N_{\text{not-}D_2} = 1, \quad N = 5. \tag{24}$$

In this case, we need, say, the Jackson Five rather than the Three Stooges (and four checks of a thousand dollars). Inserting the numbers in Eq. (24) into Eq. (20) for the likelihoods, we find

$$\Pr(A \,|\, D_2) = \frac{N_{D_2}}{N} = \frac{4}{5}, \quad \Pr(A \,|\, \text{not-}D_2) = \frac{N_{\text{not-}D_2}}{N} = \frac{1}{5}. \tag{25}$$

Inserting Eq. (18) for the priors and Eq. (25) for the likelihoods into Eq. (21) for the posterior $\Pr(D_2 \,|\, A)$, we find

$$\Pr(D_2 \,|\, A) = \frac{2}{3}, \quad \Pr(\text{not-}D_2 \,|\, A) = \frac{1}{3}. \tag{26}$$

In this case, our contestants and hence Sleeping Beauty should stay with D_2.

These results can easily be generalized from 3 to n doors (in which case we will need still more checks and Monty Hall will have to open $n-2$ doors before offering the contestant to switch). We will do this only for the case that contestants (and hence Sleeping Beauty) should be indifferent between switching and staying. In that case, we need $N = n$ candidates (and $n-1$ checks). Suppose they collectively pick the j^{th} door. We can now basically repeat the steps in Eqs. (17)–(23). The two (non-equiprobable) outcomes are:

$$D_j = \text{Checks behind } j^{\text{th}} \text{ door}, \quad \text{not-}D_j = \text{Checks not behind } j^{\text{th}} \text{ door}. \tag{27}$$

The priors are:

$$\Pr(D_j) = \frac{1}{n}, \quad \Pr(\text{not-}D_j) = \frac{n-1}{n}. \tag{28}$$

The sizes of the groups assigned to these two outcomes are:

$$N_{D_j} = n-1, \quad N_{\text{not-}D_j} = 1. \tag{29}$$

The likelihoods are:

$$\Pr(A \,|\, D_j) = \frac{N_{D_j}}{N} = \frac{n-1}{n}, \quad \Pr(A \,|\, \text{not-}D_j) = \frac{N_{\text{not-}D_j}}{N} = \frac{1}{n}. \tag{30}$$

The posterior $\Pr(D_j \,|\, A)$ is:

$$\Pr(D_j \,|\, A) = \frac{\Pr(A \,|\, D_j)\Pr(D_j)}{\Pr(A \,|\, D_j)\Pr(D_j) + \Pr(A \,|\, \text{not-}D_j)\Pr(\text{not-}D_j)}. \tag{31}$$

The posterior $\Pr(\text{not-}D_j \,|\, A)$ has the same denominator but $\Pr(A \,|\, \text{not-}D_j)\Pr(\text{not-}D_j)$ in the numerator. From Eqs. (28) and (30), we see that

$$\Pr(A \,|\, D_j)\Pr(D_j) = \Pr(A \,|\, \text{not-}D_j)\Pr(\text{not-}D_j) = \frac{n-1}{n^2}, \tag{32}$$

which means that

$$\Pr(D_j \,|\, A) = \Pr(\text{not-}D_j \,|\, A) = \frac{1}{2}. \tag{33}$$

In other words, contestants in this game (and Sleeping Beauty in the variation on the Sleeping Beauty Problem for which this game show is a proxy) should be indifferent between staying and switching.

The design of a game-show proxy for this variation on the Sleeping Beauty Problem nicely illustrates how such game shows can be used to get a better handle on such puzzles.

6 Switch or Stay? A variation on the 50:50 lifeline in "Who Wants to Be a Millionaire?"

In this section, we introduce and analyze one last variation of the game shows we used in Sections 3–5 to analyze the Sleeping Beauty Problem and generalizations of it. In those sections, the priors in Bayes's theorem were determined by symmetries of various physical systems. All else being equal, the odds of a fair coin landing heads or tails are 50/50 and one door is as likely as the other two to have the prize behind it. In the game show to be considered in this section, however, the priors will be degrees of belief that do not reflect such symmetries. In Sections 3 and 4, we showed that our analysis works regardless of whether we think of probabilities as relative frequencies or as degrees of belief. This remains true in this section. There is a simple way to represent some agent's degree of belief in some proposition A in frequentist terms. Imagine we ask our agent to place a bet (at even odds) on A or not-A. If she reports her degree of belief in A to be $0 \leq x \leq 1$, she should be prepared to spin a wheel of fortune with a pie chart printed on it with a portion $2\pi x$ corresponding to A and a portion $2\pi(1-x)$ corresponding to not-A and pick A or not-A based on where the marker of this wheel of fortune lands.

The specific game show we will examine in this section was inspired by the 50:50 lifeline of the popular TV show "Who Wants to Be a Millionaire?". Our variation on this show calls for groups of at least three contestants. Here is how it works. Regis Philbin, the host of the most successful incarnation of "Who Wants to Be a Millionaire?", asks the group of contestants some trivia question and gives them four answers to choose from. They collectively have to pick one.

Regis informs them that they have been divided into two groups ahead of time. They are told the sizes of these two groups,

$$0 < N_r < N, \quad 0 < N_w < N, \quad \text{with } N_r + N_w = N \text{ and } N_r \neq N_w \tag{34}$$

(where the subscripts 'r' and 'w' stand for 'right' and 'wrong', respectively), but not in which group they are. Regis then asks all contestants to go to one of the soundproof booths familiar from the game shows examined in Sections 3–5. If the answer they chose is right, Regis explains to the contestants before they all go into their booths, he will bring back, one by one, all those in the group with N_r contestants; if their initial answer is wrong, he will bring back, one by one, all those in the group with N_w contestants.

Every time Regis brings back a contestant, he eliminates two of the three answers not initially chosen (not necessarily the same two for every contestant) and offers the contestant a choice between the answer i initially chosen and the one remaining alternative answer not-i. As should be clear from the analysis in Section 5 (and as we will verify in detail below), a contestant will be more likely to stick to i than she otherwise would have been if $N_r > N_w$ and more likely to switch to not-i if $N_r < N_w$.

Contestants whose final answer is wrong and contestants who are not called back at all are eliminated from the game. Regis then repeats the procedure described above with a new trivia question and a reduced number of contestants. He keeps repeating it until there is only one contestant left, who is then declared the winner of the game (when it gets down to two contestants, $N_r = N_w = 1$ and the condition $N_r \neq N_w$ is dropped).

Which answer a contestant chooses when she is called back to give her final answer depends on the probabilities $\mathrm{Pr}(i)$ and $\mathrm{Pr}(\text{not-}i)$ she assigns to the answers i and not-i.[8] Unlike the doors in the Monty Hall Problem, different answers will in general be assigned different initial probabilities. Moreover, these probabilities will in general be different for different contestants. Priors and posteriors in this case clearly represent degrees of belief rather than relative frequencies, even though, as we pointed out at the beginning of this section, they can still be interpreted either way.

[8]Contestants get two pieces of evidence affecting their degree of belief in the hypothesis h that i is the right answer: e_1, which says that two answers are wrong, and e_2, which says that, depending on what the right answer is, all those in the group with N_r contestants or all those in the group with N_w contestants are called back to give their final answer.

Bayes's rule says that for some hypothesis h and two pieces of evidence e_1 and e_2 the posterior $\mathrm{Pr}(h|e_1\&e_2)$ is given by

$$\mathrm{Pr}(h|e_1\&\,e_2) = \frac{\mathrm{Pr}(e_1\&\,e_2|h)\mathrm{Pr}(h)}{\mathrm{Pr}(e_1\&\,e_2)},$$

where the prior $\mathrm{Pr}(h)$ is the degree of belief in h before the Bayesian agent found out about e_1 and e_2. The right-hand side of this equation can be rewritten either as

$$\frac{\mathrm{Pr}(e_1|h)\mathrm{Pr}(e_2|h\,\&\,e_1)\mathrm{Pr}(h)}{\mathrm{Pr}(e_1)\mathrm{Pr}(e_2|e_1)} = \frac{\mathrm{Pr}(e_2|h\,\&\,e_1)\mathrm{Pr}(h|e_1)}{\mathrm{Pr}(e_2|e_1)}$$

or as

$$\frac{\mathrm{Pr}(e_2|h)\mathrm{Pr}(e_1|h\,\&\,e_2)\mathrm{Pr}(h)}{\mathrm{Pr}(e_2)\mathrm{Pr}(e_1|e_2)} = \frac{\mathrm{Pr}(e_1|h\,\&\,e_2)\mathrm{Pr}(h|e_2)}{\mathrm{Pr}(e_1|e_2)}.$$

This shows that the order in which a Bayesian agent updates on e_1 and e_2 does not matter.

Whether this is or should also be true for a contestant in our game show evaluating the hypothesis h under consideration here in light of the two specific pieces of evidence e_1 and e_2 is a question that we leave open (one could likewise wonder whether it does or should remain true if e_1 and e_2 are the outcomes of measurements of two observables represented by non-commuting operators in quantum mechanics). We only analyze the updating on e_2 in Bayesian terms. The prior $\mathrm{Pr}(i)$ in this application of Bayes's rule is thus, in the notation of this footnote, $\mathrm{Pr}(h|e_1)$, with $h = i$.

The posterior $\Pr(i|A)$ is given by an equation similar to Eqs. (31):

$$\Pr(i \,|\, A) = \frac{\Pr(A \,|\, i) \Pr(i)}{\Pr(A \,|\, i) \Pr(i) + \Pr(A \,|\, \text{not-}i) \Pr(\text{not-}i)}, \tag{35}$$

where A is defined as "I have been called back to give my final answer aware of the protocol determining which contestants are called back to do so." Using expressions similar to those in Eqs. (30) for the likelihoods,

$$\Pr(A \,|\, i) = \frac{N_r}{N}, \qquad \Pr(A \,|\, \text{not-}i) = \frac{N_w}{N}, \tag{36}$$

and inserting $1 - \Pr(i)$ for $\Pr(\text{not-}i)$, we can rewrite the posterior as:

$$\Pr(i \,|\, A) = \frac{N_r \Pr(i)}{(N_r - N_w) \Pr(i) + N_w}. \tag{37}$$

As long as $N_r \neq N_w$ and $\Pr(i)$ is not equal to 0 or 1, the posterior differs from the prior.

We now ask what the minimum value of $\Pr(i)$ must be for a contestant to choose the answer i over the answer not-i. We are thus looking for values of the prior such that the posterior is greater than $\frac{1}{2}$. Eq. (37) tells us that $\Pr(i \,|\, A) > \frac{1}{2}$ as long as[9]

$$\Pr(i) > \frac{N_w}{N}. \tag{38}$$

The contestant should thus ask herself, factoring in that there are only two answers left but *not* how Regis selects the contestants he calls back to give their final answer, what her degree of belief $\Pr(i)$ is that i is the correct answer. If $\Pr(i) > N_w/N$, she should stay with i. If $\Pr(i) < N_w/N$, she should switch to not-i. If $\Pr(i) = N_w/N$, she should be indifferent between staying and switching.

The producers of the show can thus ensure that the selection mechanism boosts a value as low as $1/N$ for $\Pr(i)$ to a value of $\frac{1}{2}$ for $\Pr(i \,|\, A)$ by choosing $N_w = 1$ and $N_r = N - 1$. By choosing $N_r = 1$ and $N_w = N - 1$, they can likewise ensure that $\Pr(\text{not-}i) = 1/N$ gets boosted to $\Pr(\text{not-}i \,|\, A) = \frac{1}{2}$. "Who Wants to Be a Millionaire?" could thus get a new lease on life by borrowing some ideas from Sleeping Beauty and Monty Hall.

[9] Substituting $x \equiv \Pr(i)$ (with $0 < x < 1$) on the right-hand side of Eq. (37) and asking when the resulting expression is greater than $\frac{1}{2}$, we find that $N_r\, x > \frac{1}{2}\,((N_r - N_w)\, x + N_w)$ or $\frac{1}{2}(N_r + N_w)\, x > \frac{1}{2} N_w$, from which Eq. (38) follows. From equations similar to Eqs. (35)–(37), it follows that $\Pr(\text{not-}i \,|\, A) > \frac{1}{2}$ as long as $\Pr(\text{not-}i) > N_r/N$.

7 Conclusion: game shows and pedagogy

As one of us can attest on the basis of actual teaching experience (cf. note 1), the game shows introduced and analyzed in this paper make for effective pedagogical tools in introductory undergraduate courses in philosophy of science. They nicely illustrate how one can get a handle on an easy-to-grasp yet non-trivial problem by modeling it in a way that avoids one of its most confusing aspects (the problem of "self-locating beliefs", which evaporates once awakenings are replaced by contestants); by varying the parameters of the model (the solution is much easier to see for larger numbers of awakenings/contestants); and by recognizing that the problem is just one in a much broader class of similar problems (the coin flip can be replaced by rolling dice, drawing a ball from a bingo cage, spinning a roulette wheel or what have you). Our solution of the Sleeping Beauty Problem also provides an instructive example of applying Bayes's rule. In this particular application, the main challenge is to articulate exactly what one should conditionalize on when updating one's degree of belief. Using different versions of our game show, we highlighted the difference between the simple and the general form of Bayes's rule (for equal and unequal priors, respectively). We analyzed all versions of our game show both in terms of degrees of belief and in terms of relative frequencies, thus helping students appreciate the difference between these two conceptions of probability in some concrete examples. Given these pedagogical benefits, we hope that other instructors will try out these game shows in their classes. If they do, we hope that they and their students will have as much fun with them as we have had.

Acknowledgments

We are grateful to Pradeep Mutalik, whose 2016 article inspired our paper. We want to thank John Norton for helpful comments on an earlier draft, Jack Van Sambeek and Ryan Bowers for catching a serious error in Section 6, and Laurent Taudin for his lovely illustrations of our piece. We also thank Rose Adams, Jamie DeBruyckere, Terrance Gray, Scott Spicer, Charlie Heinz, and again Laurent Taudin for their contributions to the YouTube video related to this paper (see footnote 1). Finally, we thank several anonymous referees for helpful comments. The work of one of us (MJ) was supported by the Alexander von Humboldt Foundation and Jürgen Renn's Department at the Max Planck Institute for History of Science in Berlin.

References

Bradley, D. (2007). *Bayesianism And Self-Locating Beliefs.* PhD Thesis. Stanford University.

Cisewski, J., Kadane, J. B., Schervish, M. J., Seidenfeld, T., and Stern, R. (2016). Sleeping Beauty's Credences. *Philosophy of Science* 83: 324–347.

Elga, A. (2000). Self-locating belief and the Sleeping Beauty Problem. *Analysis* 60: 143–147.

Groisman, B. (2008). The End of Sleeping Beauty's Nightmare. *British Journal for the Philosophy of Science* 59: 409–416.

Hitchcock, C. (2004). Beauty and the Bets. *Synthese* 139: 405–420.

Lewis, D. (2001). Sleeping Beauty: Reply to Elga. *Analysis* 61: 171–176.

Mutalik, P. (2016). Why Sleeping Beauty Is Lost in Time. *Quanta Magazine*, March 31, 2016. Accessed April 17, 2016 at https://www.quantamagazine.org/20160331-why-sleeping-beauty-is-lost-in-time/.

Schulmann, R., Kox, A. J., Janssen, M., and Illy, J., Eds. (1998). *The Collected Papers of Albert Einstein*. Vol. 8. *The Berlin Years: Correspondence, 1914–1918*. Princeton: Princeton University Press.