Jeffrey Conditionalization: Proceed with Caution*

Forthcoming in *Philosophical Studies*

Published version: https://doi.org/10.1007/s11098-019-01356-3

Borut Trpin†

November 4, 2019

Abstract

It has been argued that if the rigidity condition is satisfied, a rational agent operating with uncertain evidence should update her subjective probabilities by Jeffrey conditionalization (JC) or else a series of bets resulting in a sure loss could be made against her (the Dynamic Dutch Book Argument). We show, however, that even if the rigidity condition is satisfied, it is not always safe to update probability distributions by JC because there exist such sequences of non-misleading uncertain observations where it may be foreseen that an agent who updates her subjective probabilities by JC will end up nearly certain that a false hypothesis is true. We analyze the features of JC that lead to this problem, specify the conditions in which it arises and respond to potential objections.

*Thanks to Kristijan Armeni, Riccardo Baratella, Mariangela Zoe Cocchiaro, Anton Donchev, Branden Fitelson, Mario Günther, Ben Levinstein, Max Pellert, Vlasta Sikimić, Reuben Stern, and especially Jan Sprenger for their comments and suggestions on earlier versions of this paper. I also want to thank for the feedback from the audiences of the 10th Arché Graduate Conference in St Andrews, UK (2017), the Workshop in Philosophy of Science at the University of Turin, Italy (2018), EENPS 2018, Bratislava, Slovakia, and Formal Epistemology Workshop in Turin, Italy (2019), where I presented the paper in its various stages. Finally, I want to thank Robbie Hopper for her help with proof-reading. I am also grateful that the research was (in parts) supported by Ernst Mach Grant, Ernst Mach Worldwide (ICM-2018-10093) and by Alexander von Humboldt Foundation.

†Munich Center for Mathematical Philosophy, LMU Munich

Contact: borut.trpin@lrz.uni-muenchen.de

1 Introduction

How should an agent update her degrees of belief when she is not fully certain of her evidence? A common prescription in Bayesian epistemology is that she needs to update by Jeffrey Conditionalization (JC), a generalization of standard Bayesian conditionalization for cases like this (see, Jeffrey, 1983, 164-183, for his explication). But why should an agent update by JC and not by some other rule? A common response is based on a proof that any agent who does not update by JC is vulnerable to a so-called dynamic Dutch book. In other words, a bookie who knows just as much as the agent can offer the agent a series of bets that the agent evaluates as fair but that lead to a guaranteed loss (Armendt, 1980). The converse was also proven: any agent who updates by JC is invulnerable to dynamic Dutch books (Skyrms, 1987).

The argument is convincing. A rational agent must avoid sure losses. However, as the problems described below show, invulnerability to Dutch books does not provide a be-all and end-all justification of JC. We will show that there exist many situations where JC gradually prescribes the agent to assign an arbitrarily high probability to a false hypothesis after observing specific sequences of uncertain but nonetheless non-misleading evidence.[1] Hence, while JC offers a pragmatic advantage (invulnerability to Dutch books), we believe that this advantage is offset by the epistemic disadvantage – a rational agent ought, after all, not assign high probability to a false hypothesis (given that the evidence is not misleading).[2] The problem is even more worrying because it is (at least in some outlined cases) robust with respect to the agent's prior probabilities. In other words, even if an agent who updates by JC is initially highly confident of the true hypothesis, there exist such sequences of non-misleading uncertain observations that she will eventually become highly confident of a false hypothesis.

2 A problematic scenario

Consider the following scenario for an illustration of how JC prescribes the agent to become highly confident of a false hypothesis: Freya is a Bayesian microbiologist. She updates her beliefs by

[1] When we say that the evidence is non-misleading, we mean that when an agent becomes more certain of some evidence E than its negation $\neg E$, E is actually the case.

[2] Although an agent cannot know whether some evidence is misleading or not in the described sense, a well-performing update rule should at least not lead to problematic outcomes in the latter cases.

Bayesian conditionalization or by JC[3] if she is not fully certain of her evidence and the rigidity condition is satisfied.[4] She has identified some bacteria in a sample and correctly believes it may only be of the A or B strain but not both. She knows that both strains have similar biochemical characteristics, except for characteristic E, which is 75% likely to be present in a given inspected part of strain A, and is present in all parts of samples containing strain B. It does not matter what her prior probability distribution is like. However, for the ease of calculations, suppose that her prior probabilities are 0.5 for both mutually exclusive and jointly exhaustive hypotheses. Further, suppose that her sample actually contains strain B, so that characteristic E is present in all inspected parts of the sample[5]. Finally, suppose Freya inspects various parts of the sample 40 times and is constantly 70% certain that she observed characteristic E in each inspected part (e.g., because her instrument only affords her ineffable learning experiences).

It is easy (if a bit lengthy) to verify that after 40 such observations Freya becomes approximately 0.99 certain that her sample contains strain A (the one where characteristic E is 0.75 likely), and merely 0.01 certain that it contains strain B which she is actually inspecting. Considering that Freya's evidence was always such that she was reasonably certain that E was present in all inspected parts of her sample (she was constantly 0.7 certain about the presence of E), it is problematic that she assigned a very high probability to strain A and a very low probability to strain B hypothesis. What went wrong in this case was that the hypothesis with the (objective) likelihood of E closest to her (subjective) certainty of observing E was favoured. But this is not what we want from an updating rule – we are not interested in confirming subjective certainties of evidence (at least when we are not fully certain). After all, Freya's observations perfectly fit strain B hypothesis as she was always more certain that E is present in the sample than that it is not.

[3] She is, after all, named after Richard Jef*frey*.

[4] The rigidity condition is satisfied when $\Pr^*(H|E_k) = \Pr(H|E_k)$ for all k, where $\Pr^*(\cdot)$ represents the posterior and $\Pr(\cdot)$ the prior probability function (Jeffrey, 1983, 174).

[5] If Freya would not always inspect different parts of her sample, she would only update her subjective probability distribution on the first inspection because her subsequent observations would contain no new evidence. I am thankful to an anonymous referee for raising this point.

Update n	$\Pr_n^*(H_A)$	$\Pr_n^*(H_B)$
Prior	0.5	0.5
1	0.6	0.4
2	0.67	0.33
3	0.72	0.28
4	0.76	0.24
5	0.80	0.20
10	0.89	0.11
20	0.96	0.04
30	0.98	0.02
40	0.99	0.01

Table 1: Select posterior subjective probabilities of hypotheses H_A (false) and H_B (true)

3 Why Jeffrey conditionalization sometimes leads astray

To understand what led to Freya's high probability of the false hypothesis we need to first formalize her probabilistic belief updating. She operated with 2 mutually exclusive and jointly exhaustive hypotheses, H_A and H_B. According to H_A, she was inspecting strain A (the presence of E is 0.75 likely in each inspected part), and according to H_B strain B (E is always present). The hypotheses were equiprobable before the first inspection because she thought it was just as likely that the bacteria would be of strain A or B (i.e., $\Pr_1(H_i) = 0.5$ for both i).[6] We assume that the presence of E in some inspected part of her sample is conditionally independent of its presence in the next (previous) inspected part given each strain. In other words, if Freya were inspecting strain A, then the conditional probability that E is present in the nth inspected part of her sample, $\Pr(E_n|H_A)$, would be 0.7 for all n.[7]

So far, so good. But why did Freya end up with a very high subjective probability of the false hypothesis (H_A)? An inspection of how her subjective probabilities evolved throughout the sequence of observations (Table 1) reveals a crucial insight: her probability of the false hypothesis H_A (the presence of E is 0.75 likely) constantly increased. Because the two hypotheses are jointly exhaustive and mutually exclusive, her probability of the true hypothesis H_B constantly decreased.

[6] Freya's priors are largely irrelevant for her subsequent high probability of the false H_A, as we show below (Theorem 1).

[7] This does not mean that the observations are probabilistically independent. Rather, the conditional probability of E_n given each hypothesis remains fixed throughout the process. If the reader finds such a conditional independence assumption unrealistic, then it should be noted that the example may be rephrased into one with a series of biased coins or dice throws instead of microbiological samples of different strains (for similar examples see, e.g., Douven, 2013; Trpin and Pellert, 2018). This conditional independence is important as $\Pr(E_n|H_i)$ is one of the key parameters in conditionalization.

Freya updated her beliefs by JC:[8,9]

$$\Pr_n^*(H_i) = \Pr_n^*(E_n)\frac{\Pr_n(H_i)\Pr(E_n|H_i)}{\Pr_n(E_n)} + \Pr_n^*(\neg E_n)\frac{\Pr_n(H_i)\Pr(\neg E_n|H_i)}{\Pr_n(\neg E_n)} \tag{JC}$$

where $0 < \Pr_n(E_n) < 1$. Note that with some algebra we obtain the equivalent form of JC (derivation omitted):

$$\Pr_n^*(H_i) = \Pr_n(H_i)\left(\Pr_n^*(E_n)\frac{\Pr(E_n|H_i) - \Pr_n(E_n)}{\Pr_n(E_n)\Pr_n(\neg E_n)} + \frac{\Pr(\neg E_n|H_i)}{\Pr_n(\neg E_n)}\right)$$
$$= \Pr_n(H_i)\mathrm{RelFact}_{n,i} \tag{JC*}$$

where $\mathrm{RelFact}_{n,i}$ represents a so-called Relative Change Factor which influences whether the posterior probability of some hypothesis increases, decreases or remains unchanged. To see why Freya's subjective probabilities of the false H_A kept increasing, we therefore need to check when $\mathrm{RelFact}_{n,i} > 1$. We obtain that $\mathrm{RelFact}_{n,i} > 1$ is equivalent to (derivation in Appendix):

$$\left(\Pr_n^*(E_n) - \Pr_n(E_n)\right)\left(\Pr(E_n|H_i) - \Pr_n(E_n)\right) > 0 \tag{1}$$

In other words, the probability of some hypothesis increases just when the agent's posterior probability (i.e., her certainty of E in that part of the sample) and the likelihood of evidence (i.e., that E is present) according to this hypothesis are both greater than or both less than her prior probability of evidence. That is, the probability of some hypothesis H_i increases when either of the following

[8] $\Pr_n(\cdot)$ represents the prior probability of some proposition before the nth update and $\Pr_n^*(\cdot)$ the posterior probability after the nth update. We represent the likelihood of evidence being present in the nth part of the sample given some hypothesis by $\Pr(E_n|H_i)$ for all i, n. Note that, given any hypothesis i, this likelihood is constant for all n because the presence of E in any part is conditionally independent of its presence in the other parts given the hypothesis under consideration. Learning about the presence of E in the nth part of the sample does therefore not affect *the conditional probability* of E being present in the $(n+1)$th part of the sample given any hypothesis. The set $\{H_i\}$, where i is either A or B, is a set of 2 mutually exclusive and jointly exhaustive hypotheses corresponding to the two strains. JC is more generally defined as $\Pr^*(H) = \sum \Pr(H|E_i)\Pr^*(E_i)$, $E_i \in E$, where E is a partition with non-zero probabilities.

[9] The rigidity condition is also satisfied. Schwan and Stern (2017) provide a convincing Causal Updating Norm (CUN) according to which rigidity is satisfied when D (a dummy variable representing the ineffable learning experience) and any arbitrary A are d-separated by an initial partition of propositions B (Schwan and Stern, 2017, 11). The causal network in Freya's example can be represented as $S \to E \to D$, where S is a variable representing the two strains which causes E_n, the presence or absence of characteristic E in the nth part, which in turn, causes D, the learning experience, because it affects whether Freya observes the characteristic in such a way that she is less than fully certain about it. CUN (and, therefore, rigidity) is satisfied because S is d-separated from D by E.

conditions (2 or 3) is satisfied:

$$\Pr_n^*(E_n) > \Pr_n(E_n) \text{ and } \Pr(E_n|H_i) > \Pr_n(E_n) \tag{2}$$

$$\Pr_n^*(E_n) < \Pr_n(E_n) \text{ and } \Pr(E_n|H_i) < \Pr_n(E_n) \tag{3}$$

Similarly, the probability of some hypothesis decreases under either of the following conditions (i.e., $\text{RelFact}_{n,i} < 1$):

$$\Pr_n^*(E_n) > \Pr_n(E_n) \text{ and } \Pr(E_n|H_i) < \Pr_n(E_n) \tag{4}$$

$$\Pr_n^*(E_n) < \Pr_n(E_n) \text{ and } \Pr(E_n|H_i) > \Pr_n(E_n) \tag{5}$$

Further, the probability of some hypothesis remains unchanged under either of the following conditions (i.e., $\text{RelFact}_{n,i} = 1$):

$$\Pr_n^*(E_n) = \Pr_n(E_n) \tag{6}$$

$$\Pr(E_n|H_i) = \Pr_n(E_n) \tag{7}$$

This suggests that JC is, in a way, all about predictions of the non-inferential evidence ($\Pr_n^*(E_n)$) in relation to its prior probability ($\Pr_n(E_n)$). The hypotheses which make predictions in line with uncertain observations become more probable and those that make wrong predictions become less probable. The hypotheses which predict no change remain just as probable as they were. If the observed evidence is as predicted, no updating takes place. Note that JC is not special in this sense – all of the above mentioned conditions (2–6) also hold for standard Bayesian conditionalization (BC), except that the conditions (3, 5) cannot be satisfied when an agent is fully certain of her evidence (because $\Pr_n(E_n)$ cannot be greater than 1). But this is a crucial distinction: BC leads an agent towards confirming what an agent is certain of, while JC goes toward confirming a subjective level of uncertainty, although we are not interested in confirming subjective levels of uncertainty but rather in confirming what is the case.

We can now show the following theorem (proof in Appendix), which explains why Freya became so highly convinced that the false hypothesis H_A was actually true (see Figure 1 for a visual

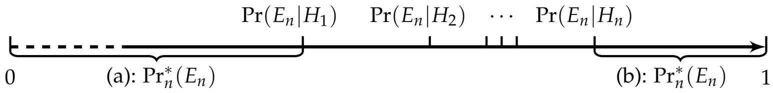

Figure 1: A representation of the situations where JC prescribes (a) an increase in the probability of H_1, or (b) an increase in the probability of H_n (given that no hypothesis is certain)[10]

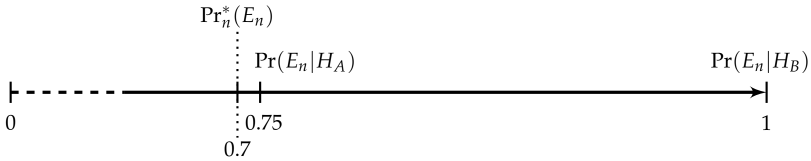

Figure 2: A representation of the three parameters that led to Freya's high confidence in the false H_A

representation of the theorem):

Theorem 1. *If the posterior probability of evidence is less (greater) than or equal to the lowest (greatest) likelihood of evidence according to some hypothesis, then JC prescribes an increase in the probability of the hypothesis according to which the likelihood of evidence is the lowest (the highest).*

Recall that Freya was always 0.7 certain that characteristic E was present in all inspected parts of her sample. Further, the likelihood of E being present in each inspected part was greater than 0.7 for both hypotheses (it was 0.75 and 1, respectively; see Figure 2). Hence, the probability of the false H_A always increased and the probability of the true H_B decreased. After sufficient belief updates the probabilities of both hypotheses converged toward 1 and 0, respectively.

Theorem 1 also shows that (if any updating would take place) Freya would assign higher probability to the true hypothesis (H_B) if she was fully certain that E was present in all inspected parts (i.e., when JC would reduce to standard Bayesian conditionalization). Her posterior probability of evidence, $\Pr^*(E_n)$, would in this case coincide with the likelihood of evidence according to H_B, $\Pr(E_n|H_B)$, and the probability of H_B, $\Pr(H_B)$ would increase.[11]

[10] Note that only the cases (a), where the posterior probability of evidence is less than or equal to the lowest likelihood of evidence, lead to problematic outcomes when the posterior probability of E is greater than that of $\neg E$. If, for example, $\Pr(E_n|H_i) < 1$ and $\Pr(E_n|H_i) \leq \Pr_n^*(E_n)$ for all i, n, then it is reasonable that the probability of the hypothesis with the largest likelihood (let us call it H_n) converges toward 1 because the hypotheses are jointly exhaustive and H_n provides the best fit for the observed sequence.

[11] JC would also not lead to a problematic outcome if the largest likelihood of evidence was less than 1, given that the posterior probability of evidence would be equal to it or greater; e.g., if she was inspecting strain B and $\Pr(E_n|H_A) = 0.7; \Pr(E_n|H_B) = 0.8$ and $\Pr_n^*(E_n) = 0.9$ for all n. In this case, $\Pr_n(H_B)$ would (accurately) converge toward 1 with increasing n. Note that Theorem 1 does not present a problem for standard Bayesian conditionalization because the posterior probability of evidence (i.e., $\Pr_n^*(E_n) = 1$) is always greater than or equal to the largest likelihood (see also footnote 10).

What is particularly important about Theorem 1 is that it applies regardless of the priors – the only parameters that lead to Freya's problem are the likelihoods and the posterior probability of evidence given each inspection. Hence, Freya could initially be highly convinced that she will observe strain B (for example, because it is much more common). However, as she always Jeffrey conditionalizes, she would nonetheless soon become confident that the strain is the false strain A. For instance, if her initial subjective probability of strain B is 0.99, it only takes her 2 updates to assign the probability of 0.47 and 0.53 to H_A and H_B, respectively. Her subsequent updates that lead to her high probability of the false hypothesis then proceed similarly as in the original example.

This is not only worrisome for our Freya in an idealized scenario. The scenario could, after all, easily be mapped to more realistic cases in which the agent is less or just as certain of her evidence than the evidence is likely according to all mutually exclusive and jointly exhaustive hypotheses. A rational agent who is in such a situation (see the case (a) in Figure 1) should, therefore, not update her subjective probabilities by JC.

4 Variations of the problematic scenario

The question that may naturally be raised is whether an agent may avoid the problems of JC when her posterior probability of evidence is greater than the likelihood of that evidence according to some, but less than the likelihood of evidence according to another hypothesis. To illustrate with Freya's example: What happens if her posterior probability of E being present in each part of her sample lies in the interval $(0.75, 1)$? Will JC still always prescribe her to update in such a way that she will subsequently become highly confident of a false hypothesis (strain A)?

The answer is negative: JC is not widely acclaimed without a reason. As we will see, however, Freya will only become more certain of the true hypothesis (strain B) if her certainty about the presence of E lies in a specific interval.

Let us show why. Suppose that everything remains the same as in the original scenario, except that Freya's certainty that E is present in each inspected part is now greater than the likelihood of E_n according to H_A and less than its likelihood according to H_B (i.e., $\Pr(E_n|H_A) = 0.75 < \Pr^*_n(E_n) < \Pr(E_n|H_A) = 1$ for all n). As it turns out, we can, again, know exactly how Freya's

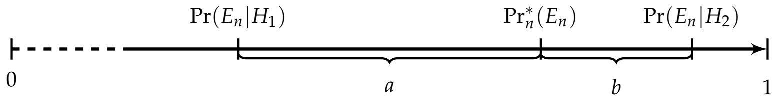

Figure 3: A representation of an instance in which JC prescribes the agent to update $\Pr^*_n(H_1)$ toward $b/(a+b)$ and $\Pr^*_n(H_2)$ toward $a/(a+b)$ by Theorem 2.

subjective probabilities in both hypotheses will end up after sufficiently many belief updates (given that $\Pr^*_n(E_n)$ is constant for all n). We can show the following theorem (proof in Appendix, see Figure 3 for a visual representation):

Theorem 2. *If the hypothesis space consists of two mutually exclusive and jointly exhaustive hypotheses and the posterior probability of different pieces of evidence with the same likelihood is constant and greater than their likelihood according to one hypothesis but less than according to the other hypothesis, then, if no hypothesis is certain, JC prescribes the agent to update in such a way that the probability of the first hypothesis converges toward $b/(a+b)$ and the probability of the second toward $a/(a+b)$, where a and b are the absolute difference between the agent's posterior probability of the pieces of evidence and their likelihood according to the first and the second hypothesis, respectively.*

In Freya's situation this means that JC will prescribe her to update in such a way that, regardless of her priors, the true hypothesis will end up as the most probable if she will always (or at least sufficiently often) be more than 0.875 certain that E is present.

To illustrate: Suppose that, for all n, the absolute difference between the likelihood of E_n on H_A and Freya's certainty that E is present in each of the inspected pieces of her sample is a, and b the absolute difference for the likelihood of E_n on H_B. The probability of H_A will then converge toward $b/(a+b)$ and the probability of H_B toward $a/(a+b)$. Because she is always more than 0.875 certain that E is present in each separate piece of her evidence, b must be less than a (see Figure 3 for a similar example). Hence, H_B will end up as the most probable hypothesis after sufficiently many updates.

It is also straight-forward to verify that the smaller b is, the closer to 1 the probability of the true H_B will converge toward. Note, however, that the probability of the true H_B will only converge toward 1 (and not just to some value close to 1) when b is 0, i.e., when the agent's posterior probability of all pieces of evidence coincides with their likelihood according to H_B, i.e., in a limiting case of Theorem 1.

5 The cases with 3 or more hypotheses

A proponent of JC may explain these problems away by objecting that situations where agents operate with merely two mutually exclusive and jointly exhaustive hypotheses are rare or even completely unrealistic. We, therefore, also need to inspect whether any general consequences of updating by JC can be predicted for cases where the agent's hypothesis space consists of 3 or more hypotheses, and especially if we can predict any problematic outcomes for these situations. As we will see, in cases of 3 or more hypotheses (unlike in the cases of only 2 live hypotheses), the priors also play a role in how subjective probabilities evolve.

Nonetheless, there are special cases where we can already predict that subjective probabilities obtained by JC will lead to problematic outcomes regardless of the priors or the number of hypotheses the agent operates with. Theorem 1, after all, holds regardless of the number of hypotheses. For an illustration, suppose that Freya is trying to determine whether the bacteria she is inspecting belongs exclusively to strain A, B, C or D, in which the presence of the biochemical characteristic E in each inspected part is 90%, 95%, 97% and 100% likely, respectively. Assume that for some reason the strain could only be one of these four. Suppose, further, that she is actually inspecting strain D and that she is always nearly 90% certain that E is present (i.e., she is highly certain about its presence). After a number of observations of different pieces of her sample JC will, again, prescribe her to become nearly certain that she is inspecting strain A, and hence nearly certain that she is not inspecting the actual one, D.

To further illustrate how concerning this outcome is, suppose that her priors for A, B, C, D are $\langle 0.01, 0.01, 0.28, 0.70 \rangle$. That is, she initially believes that she is most likely inspecting strain D (e.g., because it is the most common), although she may also be inspecting strain C, or in some exceptional cases strains A or B. JC will, nonetheless, instruct her to eventually become nearly certain that her sample contains the exceptional strain A, although her sequence of observations perfectly fits what she would expect from the most common strain D — in any case, it does not seem obvious that her inspections completely disconfirm the true hypothesis.[12]

[12] Particularly, it takes her 20 observations of E with 0.9 certainty to assign the highest probability of all to the exceptional

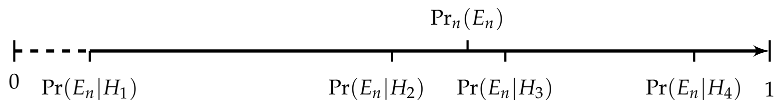

Figure 4: By Theorem 3 we can order the hypotheses according to how much their posterior probability will change in relation to their prior probability by inspecting the absolute difference between their likelihoods of evidence and the prior probability of evidence. The increasing order of relative changes in this representation is H_3, H_2, H_4, H_1.

However, not all cases are so clear cut. If the agent's posterior probability of each piece of evidence (i.e., her certainty about the evidence) lies in the open interval between the smallest and the largest likelihood of each piece of evidence and the agent is operating with more than just two hypotheses, then her prior probabilities affect how JC will prescribe her to update.

Before we can show that such scenarios may still lead to problematic outcomes, we need to first show the following theorem (proof in Appendix, see Figure 4 for a visual representation):

Theorem 3. *JC prescribes that, if any updating takes place, the probability of the hypothesis according to which the likelihood of evidence is closer to the prior probability of evidence changes less, relative to its prior probability, than other hypotheses.*

This means that, in combination with the general conditions for increased or decreased probability of a given hypothesis (Conditions 2–7), a hypothesis with a bolder prediction against the expectation of evidence receives a larger increase in its probability if the prediction is in line with the observed evidence, and a larger decrease if it is not.

Let us return to the scenario where Freya was trying to determine whether the bacteria in her sample was of the strain A, B, C or D with the likelihoods of the biochemical characteristic E in each inspected piece at 0.9, 0.95, 0.97 and 1 given each strain, accordingly. Her prior probabilities for each strain were 0.01, 0.01, 0.28 and 0.70. Her prior probability for the presence of E in the piece she first inspected was, hence, 0.9901.[13] That is, she was nearly certain that E is present in the piece she was about to inspect. Theorem 3 then implies that the posterior probability of hypothesis D will change the least in relation to its prior probability because its prediction (i.e., that E will certainly be present) is almost in line with the agent's prior probability (or expectation) of E_1.

strain A hypothesis, 27 observations to assign it a probability above 0.5, and 177 observations to assign it a very high probability (i.e., above 0.9).

[13] We calculate the prior probability of evidence by the law of total probability: $\Pr_n(E_n) = \sum_{i=1}^{m} \Pr(H_i)\Pr(E_n|H_i)$, where m is the number of hypotheses.

We need to also show the following theorem (proof in Appendix):

Theorem 4. *JC prescribes such updates that, if any updating takes place, the prior probability of the next piece of evidence E_{n+1} shifts toward the posterior probability of the previous piece of evidence E_n.*

This, again, seems reasonable. To illustrate how Theorems 3 and 4 work with Freya's example, suppose everything remains unchanged, except that she is now constantly 0.98 certain that, given each E_n, E is present in her sample. After 150 belief updates, her prior probability for E_{150} decreases to $\mathrm{Pr}_{150}(E_{150}) = 0.9806$ from her initial 0.9901 for E_1 and almost matches the level of her constant certainty that E is present in each inspected piece of her sample (0.98). Correspondingly, her probabilities for strains A, B, C, D update to $\langle 0.07, 0.02, 0.40, 0.51 \rangle$. There does not seem to be anything problematic in this outcome – she was not fully certain that E was present in her sample, so she updated her beliefs accordingly.

If her priors were, however, different, the outcome of her belief updating would also be different. For instance, suppose Freya's prior probabilities $\mathrm{Pr}_1\langle A, B, C, D \rangle$ are $\langle 0.01, 0.48, 0.48, 0.01 \rangle$. It takes her 400 updates to reach the prior probability (expectation) of E_{400} that is very close to 0.98 (particularly, 0.979), when her $\mathrm{Pr}^*_{400}\langle A, B, C, D \rangle \approx \langle 4 \times 10^{-6}, 0.06, 0.59, 0.34 \rangle$. This shows that different priors lead to different outcomes when an agent is operating with 3 or more hypotheses, but it is not particularly problematic because she was always quite but not fully certain that E was present in her sample.

6 Shifting posterior probabilities of evidence

The above described specifics of JC for cases where the hypothesis space consists of 3 or more hypotheses, however, do not mean that priors can always be set in such a way that an agent ends with reasonable subjective probabilities after a larger number of updates. As we will now show, there exist specific sequences of shifting posterior probabilities of evidence, which lead a Jeffrey conditionalizer astray regardless of her priors.

Consider the following modified version of Freya's example: she needs to detect which bacterial strain is present in her sample. For some reason, she can again only inspect whether biochemical characteristic E is present. Suppose she (correctly) knows that her sample may only contain one out of five different strains. She also knows that the strains have different tendencies for biochemical characteristic E to be present in inspected parts. The tendencies for E are as follows: the likelihood that E is present in a given piece of a sample that contains strain A is 0 (E is never present), 0.25 for strain B, 0.5 for strain C, 0.75 for strain D and 1 for strain F (E is always present).

Suppose, further, that the strain that is actually in her sample is strain F, so that characteristic E is always present. For the sake of the toy example, suppose that she believes all strains are equiprobable before she starts inspecting the sample.[14] She inspects pieces of the sample for the presence of E 250 times and her certainty that E is present in the inspected pieces shifts in such a way that she is 0.9 certain that E is present on every odd inspected piece (the first, third, fifth piece etc.) and 0.6 certain on every even inspection.

We can calculate that at the end of her inspections she becomes approximately 0.99 confident that the strain she is inspecting is strain D (and not strain F that she is actually inspecting; see Table 2). This is, again, problematic because she was, after all, always more certain that E was present than that it was not. In other words, her sequence of observations fits the strain F hypothesis perfectly, and yet she became nearly confident that her sample does not contain it. Considering she was not fully certain about her observations, we would at least not expect her to become nearly certain of a false hypothesis.

A closer inspection of Conditions 2–7 and Theorems 3 and 4 explains why Freya's sequence of uncertain observations led to such a problematic result, and, further that the priors played no role. Freya's prior probability that the first inspected piece of her sample will exhibit characteristic E was 0.5 ($\Pr_1(E_1) = 0.5$). The hypotheses H_A and H_B predicted that it was less likely for E to be present in this piece of evidence than what she expected (i.e., $\Pr(E_1|H_i) < 0.5$, where i is A or B) and the hypotheses H_D and H_F predicted it was more likely that E would be present. H_C predicted E was just as likely to be present as what she expected (i.e., H_C predicted no change in her probability of E_1).

Freya then became 0.9 certain that E was actually present in the first inspected piece. Hypotheses H_D and H_F, therefore, provided the correct predictions and their probability increased (Condition 2). H_A and H_B, however, incorrectly predicted that it would be less likely that she would

[14] As we show below, the results are independent of the priors. That is, if Freya's prior probabilities for the strains were different, it would merely take her a different number of updates by JC to assign an arbitrarily high probability to a false hypothesis.

	$\mathrm{Pr}_n^*(H_A)$	$\mathrm{Pr}_n^*(H_B)$	$\mathrm{Pr}_n^*(H_C)$	$\mathrm{Pr}_n^*(H_D)$	$\mathrm{Pr}_n^*(H_F)$	$\mathrm{Pr}_n(E_n)$	$\mathrm{Pr}_n^*(E_n)$
Pr(E_n|H_i)	0	0.25	0.5	0.75	1		
Update *n*							
0	0.2	0.2	0.2	**0.2**	**0.2**		
1	0.04	0.12	0.2	**0.28**	**0.36**	0.5	0.9
2	0.053	0.146	0.219	**0.273**	**0.309**	0.700	0.6
3	0.016	0.082	0.182	**0.300**	**0.421**	0.660	0.9
4	0.026	0.117	0.222	**0.302**	**0.334**	0.757	0.6
5	0.009	0.067	0.179	**0.316**	**0.429**	0.700	0.9
10	0.006	0.080	0.240	**0.360**	**0.314**	0.776	0.6
25	$1\text{x}10^{-4}$	0.018	0.205	**0.477**	**0.299**	0.731	0.9
50	$6\text{x}10^{-7}$	0.004	0.207	**0.629**	**0.160**	0.757	0.6
75	$8\text{x}10^{-10}$	$3\text{x}10^{-4}$	0.117	**0.750**	**0.133**	0.740	0.9
100	$5\text{x}10^{-12}$	$5\text{x}10^{-5}$	0.094	**0.835**	**0.070**	0.752	0.6
150	$5\text{x}10^{-17}$	$6\text{x}10^{-7}$	0.037	**0.935**	**0.028**	0.751	0.6
200	$7\text{x}10^{-22}$	$8\text{x}10^{-9}$	0.014	**0.976**	**0.010**	0.750	0.6
250	$9\text{x}10^{-27}$	$1\text{x}10^{-10}$	0.005	**0.991**	**0.004**	0.750	0.6

Table 2: Select posterior probabilities of hypotheses, expectations of E_n and the certainty of E's presence in a given piece of evidence

observe E, so their probability decreased (Condition 4). As a result, Freya's expectation that E is present in the next piece increased to $\mathrm{Pr}_2(E_2) = 0.7$, that is, it shifted from 0.5 for the first piece toward 0.9 for the second (Theorem 4).

We can now see why Freya eventually became very confident that the strain in her sample was strain D and not F. Her expectation that the next inspected piece will exhibit characteristic E quickly stabilized close to 0.75 (because 0.75 is the mean of the two levels of certainty that E was present in her inspected pieces, 0.6 and 0.9, toward which her expectation shifted), so her subjective probability of H_D only received small relative changes (Theorem 3). Her expectation of E in the next piece of evidence then shifted slightly over 0.75 after every odd update (because she was 0.9 certain that E is present in the inspected piece; Theorem 4) and below 0.75 after every even update (because she was only 0.6 certain that E is present). The hypothesis H_D thus always provided the correct prediction: when her expectation of E_n was slightly above 0.75 (every even round), H_D predicted that E_n was slightly less likely (exactly 0.75), just like what her subjective certainty of evidence being present suggested – she was 0.6 certain that E_n is indeed present. H_D, similarly, provided the correct predictions on every odd inspection. Her subjective probability of H_D thus slowly but surely increased (Conditions 2, 3 and Theorem 3) while her subjective probabilities of other hypotheses kept relatively increasing and decreasing in a bolder fashion (by Theorem 3).

As a matter of fact, she already ascribed the highest probability of all hypotheses to H_D after only 8 rounds and continuously after the 13$^{\text{th}}$ round. If her inspections continued and her level of certainty about E's presence continued to shift between the same levels (0.6 and 0.9), her probability of the false H_D would continue converging toward 1 by the same principles.

7 Shifting posterior probabilities of evidence: Variations

An objection could be raised that, although possible, it is highly unrealistic that an agent would undergo a sequence of uncertain observations in such a way that her certainty about the evidence would continuously shift from exactly 0.6 to 0.9 (or some other two different values). However, it is not hard to come up with more convincing scenarios of shifting posterior probabilities of evidence, where the posterior probability always shifts to a different value and yet leads the agent to a problematic endorsement of a false hypothesis. For instance, suppose that Freya is always reasonably certain that E is present in the inspected pieces of her sample, but the exact levels of her certainty constantly shift in the interval of $[0.7, 0.8]$ in a uniformly random fashion. That is, she may be 0.74 certain about E on one observation, 0.73 on the next, then 0.79, 0.77, and so on. We can predict that her expectation of E in the next inspected piece will, again, stabilize around 0.75 (by Theorem 4) and $\Pr(H_D)$ will slowly increase through a number of small increases and decreases.[15] We can thus conclude that the probability of some hypothesis according to which the likelihood of evidence is k, $k \in (0, 1)$,[16] converges toward 1 as long as the mean value of the shifting posterior probabilities of pieces of evidence is also k.

Further, updating by JC could lead to a problematic high probability for a false hypothesis even if she was inspecting another strain in which E is not always present. Suppose, for example, that Freya is trying to determine which of 11 (mutually exclusive) possible bacterial strains is present in

[15] This is because Freya's posterior probabilities of E shift in a less ordered fashion, so H_D does not always provide the correct prediction. $\Pr(H_D)$ thus also decreases but less than other hypotheses (by Theorem 3). Such uniformly random levels of evidential uncertainty, however, typically require that Freya inspects the sample 7,000 times before her probability of the false H_D reaches a very high level of 0.99. In 1,000 simulations of this scenario, she on average needed to inspect the sample 6,853 times ($\sigma = 106$) before $\Pr^*(H_D) > 0.99$.

[16] The value of k cannot be 0 or 1 because k is the mean of different (i.e., shifting) values.

her sample, with each strain exhibiting a different tendency for E in 0.1 increments from 0 to 1, and that her sample contains the strain, in which E is 0.9 likely. We can predict that JC will eventually prescribe Freya to become nearly certain that she is inspecting the strain with the 0.7 likelihoods for E (in each inspected piece of the sample) if her posterior probabilities that E is present in the sample shift in such a way that she is on average 0.75 certain about the actual presence of E and the levels of her certainty about E shift often enough.

We are able to predict this because we know that the agent will on average be 75% certain that E is present in a given inspected piece when E will actually be present (on approx. 90% of inspections) and 25% certain that E is present on the remaining 10% of inspections. Hence, we can foresee that after sufficiently many rounds she will on average be 0.7 certain that E is present in the sample ($0.9 \times 0.75 + 0.1 \times 0.25 = 0.7$) and her expectation of E will therefore converge toward 0.7. The hypothesis according to which E is 0.7 likely will on average be the closest to the expectation of E and the agent's probability of this hypothesis will slowly increase toward 1. Because the strain she is actually observing is 0.9 (and not 0.7) likely to exhibit E, the Jeffrey conditionalizer, again, ends up nearly certain that a false hypothesis is true.

More generally, an agent's subjective probability of some hypothesis will increase toward 1 in scenarios of shifting posterior probabilities of evidence if the following condition is satisfied:[17]

$$\Pr(E_n|H_i) = \frac{\sum_{m=1}^{n} \left(\Pr_m^*(E_m)\Pr(E_m|H_j) + \Pr_m^*(\neg E_m)\Pr(\neg E_m|H_j)\right)}{n} \tag{8}$$

where H_i stands for the hypothesis to which the agent ascribes an eventually increasing subjective probability, H_j for the true hypothesis (in our case, the strain that is under inspection), $\Pr_m^*(E_m)$ for the agent's posterior probability of evidence E after the mth update, and n for the number of belief updates by JC.[18]

Undergoing a sequence of observations, in which the agent's posterior probabilities of evidence shift in such a way that precisely matches the above equation is perhaps not very probable, but it nonetheless presents a problem for Jeffrey conditionalizers because the possibility of the problem-

[17] Proof omitted, although the reasoning is sketched in the previous paragraph. Note that the posterior probability of evidence needs to shift often enough.

[18] Equation 8 implies that standard Bayesian conditionalization will never lead an agent astray because it simplifies to $\Pr(E_n|H_i) = \Pr(E_n|H_j)$ when $\Pr_m^*(E_n) = 1$ for all m and $\Pr(E_m|H_j)$ is constant for all m. In other words, the hypothesis to which the agent who conditions on E_n ascribes a higher probability, H_i, is the true hypothesis H_j.

atic outcome may be foreseen.

Furthermore, the sequence of observations does not need to precisely match the conditions outlined in the above equation (8) because an approximate match may also lead the agent astray. That is, the probability for a false hypothesis will not converge toward 1, but the probability of the true hypothesis may still converge toward 0. For instance, if everything remains the same as in the case outlined in section 6 (see also Table 2), except that Freya's levels of certainty of E_n shift between 0.8 and 0.6 (instead of 0.9 and 0.6), then the probability of the true hypothesis H_F still converges toward 0, while the probabilities of H_C and H_D (both false) converge toward 0.19 and 0.81, respectively.

8 Misleading evidence and theory-laden observations

There are two potential objections that may be raised against the implications of JC as presented in the previous sections. It may be argued that:

1. The evidence the agent learned in the examples was (*contra* Fn. 1) either actually misleading, or else the outcome prescribed by JC was rationally required, and
2. Observations for a Bayesian are (or rationally should be) theory-laden, which would block problematic belief updating in the described examples.[19]

Both potential objections are related: the first provides the starting point for a discussion of subjective and objective aspects of evidence, while the second provides one potential resolution of the diagnosed problems of JC.

8.1 Subjective and objective aspects of evidence

What Freya's examples demonstrate is that JC adapts a subjective probability distribution in such a way that it supports the uncertain learning experience and not primarily the evidence itself. In other words, when Freya's observations suggest that E is more likely present than what she anticipated, she becomes more confident about the hypotheses on which E is more likely and less

[19] Thanks to an anonymous referee for this journal and an anonymous referee for the Formal Epistemology Workshop 2019 for bringing these objections to my attention.

confident about those on which E is less likely. As we saw, this may lead to situations in which experience suggests E is more likely than $\neg E$ and yet the agent becomes more confident about the hypothesis which supports $\neg E$.

One might argue that this is not really a problem as these learning experiences are actually misleading and only intuitively seem not to be such. To simplify the discussion, we will rephrase Freya's situation into one where an agent needs to determine what die from a collection of dice with various configurations of red or blue faces is being thrown. That is, instead of referring to presence of a characteristic E in various pieces of a microbiological sample, we can simply refer to a game in which an agent needs to identify the correct die.

Suppose your friend has two dice: one with all 6 faces in blue (6B) and another with 2 red and 4 blue faces (2R4B). Unbeknownst to you, your friend selects the die with all faces in blue and repeatedly rolls it. You observe the outcomes in dim light, so you are not completely sure what color you see when you check the die, but it seems to you that it is most likely blue, and if it is not blue, then it can only be red (as you know which two dice your friend selected from).

To make the example more precise: suppose that prior to the roll of the die you assign equal probability to each die. Hence, your prior probability for the die landing red is (by the law of total probability) $\frac{1}{6} \approx 0.17$. Your observations of the die are such that you repeatedly become 0.7 certain that it landed blue and 0.3 that it may have landed red. This suggests that it is more likely that the die landed red than you expected prior to the observations (initially, for example, 0.3 vs. 0.17). Hence, the argument goes, you should increase your confidence that your friend chose the die with 2 red and 4 blue faces (and not the one with 6 blue faces) as this hypothesis supports your learning experience – it increases your confidence in red. This is also what updating by JC leads to. As you end up assigning very low probability to the true and very high probability to the false hypothesis after a number of observations, your evidence has to actually be misleading.

It needs to be stressed that just because your observations suggest that the die may have repeatedly also landed red, this on its own does not suffice for ruling out the hypothesis according to which the die only has blue faces. By Theorems 1 and 2 we know that you will only end up more than 50% confident in the false hypothesis (2 red, 4 blue faces) if you are often enough more than $\frac{1}{6} \approx 0.17$ certain that the die landed red. Further, the probability of the true hypothesis (all blue faces) will converge toward 0 only if you are often enough more than or exactly $\frac{2}{6} \approx 0.33$ certain that the die landed red. In the above example (the die always lands blue and you are 0.7 certain of blue and 0.3 of red) your probability for the false die with 2 red faces will converge toward 0.9 (by Theorem 2), and not toward 1 (as in Freya's examples).

Nevertheless, the argument continues, there is nothing problematic in this outcome, neither is there anything problematic in analogous Freya's examples: the experiences raise your credence in red (and hence lower your credence in blue), so you need to adapt your credences in the rest of the propositions accordingly. In other words, when experience supports red, you need to become more confident about hypotheses on which red is more likely. If it turns out that you end up increasingly more confident about a false hypothesis, it is not the updating rule that is to blame but rather the evidence that is misleading.

There are at least two responses to this potential objection. First, it highlights the difference between objective and subjective aspects of evidence under uncertainty. Second, it nicely illustrates what JC does in practice: it primarily supports subjective aspects of a learning experience and not of the evidence itself.

We have mentioned (Fn. 1) that evidence of some "binary"[20] E is misleading when you become more than 0.5 confident of E, although E is not the case. According to this view, the evidence is misleading if a die that may only land blue or red lands blue, but your experience suggests that it is more likely that it landed red (i.e., if $\Pr^*(red) > 0.5, \Pr^*(blue) < 0.5$). In the case with the die, the evidence is therefore not misleading in this sense: the die is the one with all blue faces and you are constantly 0.7 certain that it landed blue and 0.3 certain that it landed red. The same can also be said of all Freya's examples in the previous sections. The learning experiences suggest that the true outcome was the most likely and are, in this sense, not misleading.

The definition of misleading evidence according to the potential objection is different: on this view the evidence is misleading if you end up further from the truth by assigning lower probability to the true hypothesis (as measured, for instance, by Brier's rule; Brier, 1950). The evidence is misleading if your subjective learning experience goes against your prior probability of evidence and thus points you in the wrong direction. That is, when you assign equal probability to the

[20] "Binary" E here represents a simple partition into E and $\neg E$, the only kind of evidential partitions we deal with in this paper.

two hypotheses about the die (all blue, and 4 blue, 2 red faces), your prior probability for blue is 0.83. As the die your friend selected only has blue faces, any experience that lowers your credence in blue (that is, any learning experience where $\Pr^*_n(blue) < 0.83$) represents an experience with misleading evidence. The examples discussed here are therefore operating with what is in this sense considered misleading evidence.

This highlights what we identify as the main problematic aspect of JC: it pays too much attention to subjective aspects of evidence. Clearly, any learning experience that leads an agent further from the truth (in the sense of becoming less confident about true propositions) is misleading from the subjective point of view of the agent. But this does not mean that uncertain evidence itself is misleading in a more objective sense. Why should agent's subjective probabilities determine whether uncertain evidence is misleading? Furthermore, updating by JC also leads to unsatisfactory outcomes in certain cases where the evidence is not misleading in this alternative, more subjective sense.

By way of illustration, let us consider two cases where the die always lands blue and you are 0.7 certain that it landed blue and 0.3 that it landed red. In the first case you initially assign equal (0.5) probability to both live hypotheses (4 red, 2 blue, and all blue faces). In the second, you start with 0.99 and 0.01 probabilities for the two hypotheses. Updating by JC in the first case will lead you to slowly decrease the probability of the true hypothesis (all blue) from 0.5 towards 0.1. In the latter case, however, you will slowly increase the probability of the true hypothesis from 0.01 towards 0.1.

This is because your probability for red and blue will eventually exactly match your observations in both cases, which can only happen if $\Pr(2R4B)$ updates to 0.9 (by Theorem 4[21]). You end up assigning very low probability to the true hypothesis in both cases. The only difference is that in the first case you decrease your confidence in the true hypothesis, while in the second case you increase it until your confidence in red and blue matches your learning experiences. As the example demonstrates, this may still mean that you will end up assigning very low probability to the true hypothesis (in this case, 0.1 to 6B), even if your observations and updates by JC initially support increased confidence in the true hypothesis.

The problem does therefore not lie in the fact that the evidence decreases your confidence in

[21] $\Pr(red) = \Pr(2R4B)\Pr(red|2R4B) + \Pr(6B)\Pr(red|6B) = 0.9 \times \frac{1}{3} + 0.1 \times 0 = 0.3 = \Pr^*(red)$

what is the case (here, that the die landed blue), but rather that JC prescribes you to update your credences in such a way that they support your subjective learning experience (here, that you were 0.7 certain that the die landed blue and 0.3 that it landed red).

We believe that this highlights a crucial aspect of JC that may make the rule perfectly applicable for some cases but not for all (certainly not for those outlined in this paper): it gives too much weight to subjective aspects of learning experiences – how certain you were about the evidence – while not paying enough attention to "objective" aspects of the evidence, that is, whether some evidence E really is the case or not.

This also further demonstrates why standard Bayesian conditionalisation (i.e., updating credences after learning $\Pr_n^*(E) = 1$) does not suffer from the diagnosed problems: the fact that newly learned evidence is learned with full certainty means that the agent will update her credences in such a way that they will support what actually is the case.[22]

One potential solution in cases like this is that instead of updating by JC we update our credences by following a probabilistic rule that gives additional weight to evidence instead of accounting just for the subjective aspects of the learning experience. One such rule that was proposed in the literature is, for instance, a probabilistic form of inference to the best explanation of uncertain evidence that may outperform JC in some cases (see, e.g., Trpin and Pellert, 2018). Following such an alternative rule presents a potential resolution of these problems, but it comes with a price: by not updating by JC we expose ourselves to potential Dutch books (Skyrms, 1987) and incoherence (Climenhaga, 2017).

8.2 Should observations be theory-laden?

This discussion of subjective and objective aspects of learning experiences also provides grounds for another related response. According to this potential objection we can keep using JC in uncertain evidential situations by restricting its use to situations where, intuitively speaking, agents consider that their learning experiences contain significant evidence.[23] The problems discussed above would thus be resolved: we would still be able to use JC when it is appropriate to pay atten-

[22] We are assuming that agents only become fully certain if something is actually the case. See, however, Rescorla (2019) for an interesting discussion of non-factive aspects of conditionalisation.

[23] The term "significant evidence" here simply means substantial or important evidence. It is in no way related to statistical significance or low p-values.

tion to subjective aspects of a learning experience. However, we would avoid apparent pitfalls of JC in cases where it may be foreseen that it may lead an agent astray.

In other words, one may argue that observations for Bayesian agents are (or should be), in a sense, theory-laden. This means that whether a rational agent ought to update by JC should not depend only on the rigidity condition (as requested by Jeffrey, 1983, 174), but also on other theoretical considerations of the situation, e.g., whether evidence is considered significant, which may depend on agent's prior probability distribution.

It is best to provide an example to illustrate the suggestion. Suppose that three agents are trying to determine which die their friend is throwing. For whatever reason, agent A is 0.999 confident that the die has all faces in red (6R) and if not, then all in blue (6B). Agent B's credences are just the opposite: she assigns 0.001 confidence to 6R and 0.999 to 6B. Agent C, on the other hand, assigns 0.999 probability to the hypothesis that the die has 3 red and 3 blue faces (3R3B) and if not, then 6R and 6B are very unlikely but equiprobable.

Now suppose that the die is thrown and the agents are all uncertain what face it landed on, for instance because they observe the outcome in dim lightning or perhaps they are only able to inspect the outcome on a murky photograph of the outcome. Importantly, all agents have phenomenologically the same experience when they observe the outcome: they are all 0.5 certain that the die landed on a red face and 0.5 that it did not (i.e., that it landed on blue).

Updating by JC would mean that agents A and B end up with the same posterior probabilities assigned to hypotheses 6B and 6R (i.e., $\Pr^*(6B) = \Pr^*(6R) = 0.5$), while agent C's credences would remain unchanged (by Condition 6). As we noted earlier, their credences update in such a way that they support the learning experience (that is, they accommodate $\Pr^*(R) = 0.5$; see also Table 3).

However, it may be argued that such an outcome is unconvincing because we are assuming that all agents should respond in the same way to a phenomenologically same learning experience. In assuming so we commit ourselves to what we may call the Independence of Observations Assumption (IOA):[24]

[24] Thanks to an anonymous referee for this journal for pointing out this assumption.

$\Pr(\cdot)$	6B	3B3R	6R	R		$\Pr^*(\cdot)$	6B	3B3R	6R	R
A	0.999	0	0.001	0.001			0.5	0	0.5	0.5
B	0.001	0	0.999	0.999	$\xrightarrow[\Pr^*(B)=0.5]{\Pr^*(R)=0.5}$		0.5	0	0.5	0.5
C	0.0005	0.999	0.0005	0.5			0.0005	0.999	0.0005	0.5

Table 3: Changes in credences of agents A, B, and C after an update by JC

IOA: When rational Bayesian agents make phenomenologically the same observation of $E/\neg E$, their subjective probabilities to $E/\neg E$ after the observation should be same regardless of their prior probability assignments to $E/\neg E$.

The example with the three agents suggests that IOA does not hold or at least that it should not hold for rational agents. Considering that the evidence was highly uncertain (it was as likely that the die landed red as blue) it seems that agents A and B should simply ignore the learning experience and stick to their guns: they should (just like agent C) not change their credences at all, which would happen if the observation was interpreted as showing red with 0.001 (for agent A) or 0.999 probability (for agent B). This is a particularly convincing suggestion because agents A and B were initially highly confident about one of the extreme hypotheses (6B/6R), yet they end up assigning equal probability to both live hypotheses after a highly uncertain observation. A resolution of the problem could therefore be to simply reject IOA and omit updating credences in situations like that.

On the other hand, if rational agents nevertheless update in line with JC (ending up in a state shown in Table 3), then it could be said that they consider the learning experience to be of high significance. It may be argued that observations for Bayesians need to be theory-laden in this sense: they need to consider whether they are willing to allow the uncertain learning experience to shift their subjective probability distribution. One of the simplest criteria for deciding significance of evidence could be prior probabilities. If an agent's prior probability for some hypothesis is, for instance, above a contextually defined threshold (e.g., 0.99) and uncertain learning experience would substantially lower it (e.g., from 0.999 to 0.5), then a rational agent needs to consider whether the learning experience is significant enough for a substantial shift of her credences or not. In other words, an agent needs to proceed with caution when updating by JC.

It is hard to disagree with the suggestion that we should give up on IOA and require JC to not be a mindless belief updating rule that is applied in every uncertain evidential situation where rigidity condition is satisfied. Even the rigidity condition is, in a sense, mindless (it is required by probability calculus). Moreover, the suggestion also resolves the problematic outcomes in Freya's cases. If Freya's learning experience suggests E is more likely than $\neg E$, then she may simply avoid JC when updating by that rule would reduce her probability for E. She is still allowed to update by JC, but only if she thinks that her learning experience contains substantial evidence (i.e., substantial according to her other theoretical considerations). Finally, this suggestion allows us to retain JC as a rational belief updating rule despite the potentially problematic outcomes. We need a rule that governs belief updating in uncertain evidential situations as we do not have a fully convincing alternative.

Before we fully embrace the suggestion that IOA needs to be rejected in favor of JC that is based on agent's theory-laden observations, it needs to be pointed out that the suggestion comes with a couple of downsides. First and foremost, it bolts a philosophical consideration onto JC which is, in principle, a simple theorem that extends standard Bayesian conditionalisation when the rigidity condition is satisfied.[25] The discussed scenarios, however, suggest that further (philosophically motivated) restrictions to applications of JC should be made, so perhaps this is not problematic. It should be noted, though, that in contrast to the rigidity condition rejecting IOA could make the rule too flexible. We want rational rules of belief updating to be normative and provide guidance when dealing with (uncertain) evidence. The flexibility of JC that arises after rejecting IOA could make the rule, in a sense, unfalsifiable: if the outcome of (repeated) updating by JC is as expected, then it is business as usual. If not, then the agent should either be more cautious and not update by JC or the outcome was expected if uncertain evidence is taken seriously.

Nevertheless, how to proceed when IOA is rejected and observations are required to be theory-laden could be spelled out with more precision in future work. It seems, however, that the new guiding principle which determines whether uncertain evidence is substantial or not should not rely (just) on the priors. Recall the example where an agent starts with 0.99 probability for the hypothesis that a die has 4 blue and 2 red faces (4B2R) and 0.01 for 6B, and then observes blue

[25] Suppose $\Pr(H|E) = \Pr^*(H|E)$ and $\Pr(H|\neg E) = \Pr^*(H|\neg E)$, i.e. the rigidity condition. It is then trivial to show that $\Pr^*(H) = \Pr(H|E)\Pr^*(E) + \Pr(H|\neg E)\Pr^*(\neg E)$, i.e. the rule of Jeffrey conditionalization for binary E (e.g. Jeffrey, 1983, 169).

with 0.7 certainty (and red with 0.3). Her belief updating could easily be blocked if the decision to update by JC or not would depend on the priors. The observations (probability of red is 0.3) decrease the probability of the strongest hypothesis ($\Pr(4\text{B}2\text{R}) = 0.99$). This should be good news as the strongest hypothesis is in fact false. In other words, highly inaccurate prior probabilities may prevent us from ever improving our epistemic state if IOA is rejected and the decision to update by JC or not relies on how it would affect strong hypotheses. The probability of the true hypothesis should ideally converge toward 1 (or at least increase) regardless of the priors.

Rejecting IOA therefore presents one possible resolution of these problems but it needs to be further analyzed and formulated more precisely before we can fully embrace it. Nevertheless, we believe that this is a promising idea that is worth exploring in future research. After all, applications of probability to epistemological cases could benefit from the art of judgement (to paraphrase Jeffrey, 1992).

9 Conclusion

It is worth noting that the problems reported here typically only appear when a longer sequence of updates by JC is considered. When we inspect single updates (e.g., the first 5 updates reported in Table 2), the updating rule does not appear to be problematic at all, especially if there are more than just 2 or 3 hypotheses in the agent's hypothesis space.

Further, the cases where the levels of agent's evidential certainty are shifting throughout a sequence of observations are the most similar to real-life situations of those we considered. As we demonstrated, these cases do not always lead to problematic outcomes. The sheer possibility that they may, however, suggests that a rational agent should either proceed with caution before updating her beliefs by JC or avoid it completely because JC prescribes belief updates are done in such a way that the levels of subjective certainty of evidence are confirmed. For instance, if one is *somewhat certain* of some evidence E, then the hypothesis according to which E is *somewhat likely* will increase in probability.

This is not what a rational agent who is updating her beliefs in uncertain evidential situations should be interested in. On the contrary, she should be interested in the uncertain evidence itself, not her subjective levels of certainty (at least not in the confirmatory sense). In Freya's cases the agent was actually interested in whether the characteristic E was or was not present in the inspected parts of the sample because this was the key to correct bacterial strain identification. Similarly, in the cases with two-colored dice, the agents should be interested in which die they are looking at and not primarily how certain they were about the observations. It therefore remains an open question for further research how a rational agent needs to update beliefs in uncertain evidential situations: be it by an adaptation of JC (e.g., by precisely specifying when it should be omitted), an alternative rule, or a contextually-dependent pluralistic rule that applies JC in some and alternative rules in the other cases.

References

Armendt, B. (1980). Is there a Dutch book argument for probability kinematics? *Philosophy of Science 47*(4), 583–588.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review 78*(1), 1–3.

Climenhaga, N. (2017). Inference to the best explanation made incoherent. *Journal of Philosophy 114*(5), 251–273.

Douven, I. (2013). Inference to the best explanation, Dutch books, and inaccuracy minimisation. *Philosophical Quarterly 63*(252), 428–444.

Jeffrey, R. (1983). *The Logic of Decision*. Chicago and London: University of Chicago Press.

Jeffrey, R. (1992). *Probability and the Art of Judgment*. Cambridge University Press.

Rescorla, M. (2019, Jun). On the proper formulation of conditionalization. *Synthese*.

Schwan, B. and R. Stern (2017). A causal understanding of when and when not to Jeffrey conditionalize. *Philosophers' Imprint 17*(8), 1–21.

Skyrms, B. (1987). Dynamic coherence and probability kinematics. *Philosophy of Science 54*(1), 1–20.

Trpin, B. and M. Pellert (2018). Inference to the best explanation in uncertain evidential situations. *The British Journal for the Philosophy of Science*.

A Appendix: Proofs

A.1 Derivation of Inequality 1

First note that $\text{RelFact}_{n,i} > 1$ is equivalent to:

$$\left(\Pr_n^*(E_n)\frac{\Pr(E_n|H_i) - \Pr_n(E_n)}{\Pr_n(E_n)\Pr_n(\neg E_n)} + \frac{\Pr(\neg E_n|H_i)}{\Pr_n(\neg E_n)}\right)$$
$$= \frac{\Pr_n^*(E_n)\Pr(E_n|H_i) - \Pr_n^*(E_n)\Pr_n(E_n) + \Pr(\neg E_n|H_i)\Pr_n(E_n)}{\Pr_n(E_n)\Pr_n(\neg E_n)} > 1 \quad (9)$$

where $0 < \Pr_n(E_n) < 1$. Because $\Pr_n(E_n)\Pr_n(\neg E_n) > 0$, we derive:

$$\Pr_n^*(E_n)\Pr(E_n|H_i) - \Pr_n^*(E_n)\Pr_n(E_n) + \Pr(\neg E_n|H_i)\Pr_n(E_n) > \Pr_n(E_n)\Pr_n(\neg E_n) \quad (10)$$

After substituting $\Pr(\neg E_n|H_i)$ with $1 - \Pr(E_n|H_i)$ and $\Pr_n(\neg E_n)$ with $1 - \Pr_n(E_n)$ and rearranging the Inequality 10, we obtain:

$$\Pr_n^*(E_n)\Pr_n(E_n|H_i) - \Pr_n^*(E_n)\Pr_n(E_n) - \Pr_n(E_n|H_i)\Pr_n(E_n) + \big(\Pr_n(E_n)\big)^2 > 0 \quad (11)$$

It is trivial to see that Inequality 11 is an expanded form of Inequality 1. This concludes our derivation.

Note that by following the same procedure, we also derive that $\text{RelFact}_{n,i} < 1$ is equivalent to:

$$\big(\Pr_n^*(E_n) - \Pr_n(E_n)\big)\big(\Pr(E_n|H_i) - \Pr_n(E_n)\big) < 0 \quad (12)$$

Similarly, $\text{RelFact}_{n,i} = 1$ is equivalent to:

$$\big(\Pr_n^*(E_n) - \Pr_n(E_n)\big)\big(\Pr(E_n|H_i) - \Pr_n(E_n)\big) = 0 \quad (13)$$

A.2 Proof of Theorem 1

Theorem 1. *If the posterior probability of evidence is less (greater) than or equal to the lowest (greatest) likelihood of evidence according to some hypothesis, then JC prescribes an increase in the probability of the* *hypothesis according to which the likelihood of evidence is the lowest (the highest).*

Proof. Suppose that the likelihood of evidence is the lowest according to hypothesis H_k and the highest according to H_q. We first observe the following simple consequence of the law of total probability (assuming no hypothesis is certain):

$$\Pr(E_n|H_k) < \Pr_n(E_n) < \Pr(E_n|H_q) \tag{14}$$

Hence, if the posterior probability of evidence is less than or equal to the lowest likelihood of evidence, then the posterior probability of evidence is also less than the prior probability of evidence:

$$\Pr_n^*(E_n) \leq \Pr(E_n|H_k) < \Pr_n(E_n) \tag{15}$$

Recall that, for any i, the probability of H_i increases if the following condition is satisfied:

$$\Pr_n^*(E_n) < \Pr_n(E_n) \text{ and } \Pr(E_n|H_i) < \Pr_n(E_n) \tag{3}$$

It then follows from Condition 3 and Inequality 15 that $\Pr_n^*(H_k) > \Pr_n(H_k)$ when $\Pr_n^*(E_n) \leq \Pr(E_n|H_k)$.

Similarly, if the posterior probability of evidence is greater than or equal to the highest likelihood of evidence, it is also greater than the prior probability of evidence:

$$\Pr_n(E_n) < \Pr(E_n|H_n) \leq \Pr_n^*(E_n) \tag{16}$$

Recall that, for any i, the probability of H_i increases if the following condition is satisfied:

$$\Pr_n^*(E_n) > \Pr_n(E_n) \text{ and } \Pr(E_n|H_i) > \Pr_n(E_n) \tag{2}$$

It then follows from Condition 2 and Inequalitiy 16 that $\Pr_n^*(H_q) > \Pr_n(H_q)$ when $\Pr_n^*(E_n) \geq \Pr(E_n|H_q)$. This concludes our proof of Theorem 1. □

A.3 Proof of Theorem 2

Theorem 2. *If the hypothesis space consists of two mutually exclusive and jointly exhaustive hypotheses and the posterior probability of different pieces of evidence with the same likelihood is constant and greater than their likelihood according to one hypothesis but less than according to the other hypothesis, then, if no hypothesis is certain, JC prescribes the agent to update in such a way that the probability of the first hypothesis converges toward* $b/(a+b)$ *and the probability of the second toward* $a/(a+b)$*, where* a *and* b *are the absolute difference between the agent's posterior probability of the pieces of evidence and their likelihood according to the first and the second hypothesis, respectively.*

Proof. Suppose that the posterior probability of the nth piece of evidence is greater than the likelihood of evidence according to H_k and less than the likelihood of evidence according to H_q:

$$\Pr(E_n|H_k) < \Pr_n^*(E_n) < \Pr(E_n|H_q) \tag{17}$$

Note that we can represent the likelihood $\Pr(E_n|H_k)$ as $\Pr_n^*(E_n) - a, a > 0$ and the likelihood $\Pr(E_n|H_q)$ as $\Pr_n^*(E_n) + b, b > 0$. Additionally, note that $\Pr_n(H_q) = 1 - \Pr_n(H_k)$. Finally, note that because H_k and H_q are mutually exclusive and jointly exhaustive, $\Pr_n(E_n) = \Pr(E_n|H_k)\Pr_n(H_k) + \Pr(E_n|H_q)\Pr_n(H_q)$.
Suppose $\Pr_n^*(E_n) < \Pr_n(E_n)$. Hence:

$$\begin{aligned}\Pr_n^*(E_n) - \Pr_n(E_n) &= \Pr_n^*(E_n) - \Pr_n(H_k)\left(\Pr_n^*(E_n) - a\right) - \left(1 - \Pr_n(H_k)\right)\left(\Pr_n^*(E_n) + b\right) \\ &= \Pr_n(H_k)\,(a+b) - b < 0 \end{aligned} \tag{18}$$

Because $a + b > 0$, it follows that:

$$\Pr_n(H_k) < b/(a+b) \tag{19}$$

We have now shown that $\Pr(E_n|H_k) < \Pr_n^*(E_n) < \Pr_n(E_n)$ implies that $\Pr_n(H_k) < b/(a+b)$. Recall, again, that by Condition 3, $\Pr(E_n|H_i) < \Pr_n(E_n)$ and $\Pr_n^*(E_n) < \Pr_n(E_n)$ imply $\Pr_n^*(H_i) > \Pr_n(H_i)$ for any i. We thus conclude that $\Pr(E_n|H_k) < \Pr_n^*(E_n) < \Pr_n(E_n)$ also implies $\Pr_n^*(H_k) > \Pr_n(H_k)$. The probability of H_k, therefore, increases if it is less than $b/(a+b)$.

It is trivial to see that $\Pr(E_n|H_k) < \Pr_n(E_n) < \Pr_n^*(E_n)$, on the other hand, implies $\Pr_n(H_k) >$

$b/(a+b)$. Recall that by Condition 4, $\Pr(E_n|H_i) < \Pr_n(E_n)$ and $\Pr^*_n(E_n) > \Pr_n(E_n)$ imply $\Pr^*_n(H_i) < \Pr_n(H_i)$ for any i. We conclude that $\Pr(E_n|H_k) < \Pr_n(E_n) < \Pr^*_n(E_n)$ also implies $\Pr^*_n(H_k) < \Pr_n(H_k)$. The probability of H_k, therefore, decreases if it is greater than $b/(a+b)$. Because we already know that the probability of H_k increases if it is less than $b/(a+b)$, we conclude that it converges toward $b/(a+b)$.

To finish our proof we need to show that the probability of H_q converges toward $a/(a+b)$. Because $\Pr_n(H_q) = 1 - \Pr_n(H_k)$, it immediately follows that the probability of H_q converges toward $1 - (b/(a+b)) = a/(a+b)$. Because a and b are constant, this concludes our proof of Theorem 2.

□

A.4 Proof of Theorem 3

Theorem 3. *JC prescribes that, if any updating takes place, the probability of the hypothesis according to which the likelihood of evidence is closer to the prior probability of evidence changes less, relative to its prior probability, than other hypotheses.*

Proof. We need to inspect the absolute difference $|\text{RelFact}_{n,i} - 1|$ to see how much $\Pr^*_n(H_i)$ changes in relation to its prior probability $\Pr_n(H_i)$ (regardless of direction). We first obtain:

$$|\text{RelFact}_{n,i} - 1| = \left| \frac{(\Pr^*_n(E_n) - \Pr_n(E_n))\,(\Pr_n(E_n|H_i) - \Pr_n(E_n))}{\Pr_n(E_n)\Pr_n(\neg E_n)} \right| \tag{20}$$

We can now compare which of the two hypotheses (H_i or H_j) relatively changes more in the same round. By applying the rules of absolute products and absolute quotients, we obtain that $|\text{RelFact}_{n,i} - 1| < |\text{RelFact}_{n,j} - 1|$ is equivalent to:

$$|\Pr(E_n|H_i) - \Pr_n(E_n)| < |\Pr(E_n|H_j) - \Pr_n(E_n)|; i \neq j \tag{21}$$

This means that the posterior probability of the hypothesis according to which the likelihood of evidence is closer to the prior probability of evidence will change less relative to its prior probability. This concludes our proof. □

A.5 Proof of Theorem 4

Theorem 4. *JC prescribes such updates that, if any updating takes place, the prior probability of the next piece of evidence E_{n+1} shifts toward the posterior probability of the previous piece of evidence E_n.*

Proof. We know that no updating takes place if $\Pr^*_n(E_n) = \Pr_n(E_n)$ (see Equation 13). Hence, we need to show that when $\Pr^*_n(E_n)$ is greater (less) than $\Pr_n(E_n)$, $\Pr_{n+1}(E_{n+1})$ increases (decreases) at most to the level of $\Pr^*_n(E_n)$:

$$\Pr_n(E_n) < \Pr^*_n(E_n) \Rightarrow \Pr_n(E_n) < \Pr_{n+1}(E_{n+1}) \leq \Pr^*_n(E_n) \tag{22}$$

$$\Pr_n(E_n) > \Pr^*_n(E_n) \Rightarrow \Pr_n(E_n) > \Pr_{n+1}(E_{n+1}) \geq \Pr^*_n(E_n) \tag{23}$$

Let us start by proving that (22) holds. Assume $\Pr_n(E_n) < \Pr^*_n(E_n)$. We will first show that it follows that $\Pr_n(E_n) < \Pr_{n+1}(E_{n+1})$.

We know that when the posterior probability of evidence is greater than the prior probability of evidence (our assumption), the posterior probability of hypotheses according to which the evidence is more likely than its prior probability increases (Inequality 1) and, similarly, the posterior probability of hypotheses according to which the evidence is less likely than its prior probability decreases (Inequality 12). Suppose the hypotheses are ordered in accordance to the likelihood of evidence with the hypothesis with the lowest likelihood being H_1 and the hypothesis with the highest likelihood H_m. Further, suppose the likelihood of evidence according to hypotheses H_1 to H_{n-1} is less than its prior probability in round n, equal to the prior probability of evidence according to H_n, and greater than the prior probability of evidence for hypotheses H_{n+1} to H_m. Recall also that for all i, n, $\Pr(E_{n+1}|H_i) = \Pr(E_n|H_i)$ (the assumption of independence of various pieces of evidence). Hence, the prior probability of the next piece of evidence (E_{n+1}) in round $n + 1$ updates to:

$$\begin{aligned}\Pr_{n+1}(E_{n+1}) = \big(\Pr_n(H_1) - a\big)\Pr(E_n|H_1) + \ldots + \big(\Pr_n(H_{n-1}) - b\big)\Pr(E_n|H_{n-1})\\ + \Pr_n(H_n)\Pr(E_n|H_n) + \big(\Pr_n(H_{n+1}) + c\big)\Pr(E_n|H_{n+1}) + \ldots + \big(\Pr_n(H_m) + d\big)\Pr(E_n|H_m)\end{aligned} \tag{24}$$

where $a, \ldots, b, c, \ldots, d$ are all greater than 0. We can simplify (24) to:

$$\mathrm{Pr}_{n+1}(E_{n+1}) = \mathrm{Pr}_n(E_n) - a\mathrm{Pr}(E_n|H_1) - \ldots - b\mathrm{Pr}(E_n|H_{n-1}) + c\mathrm{Pr}(E_n|H_{n+1}) + \ldots + d\mathrm{Pr}(E_n|H_m) \tag{25}$$

We can now show that $\mathrm{Pr}_{n+1}(E_{n+1}) > \mathrm{Pr}_n(E_n)$. Assume the following inequality for reductio ad absurdum:

$$\mathrm{Pr}_{n+1}(E_{n+1}) \leq \mathrm{Pr}_n(E_n) \tag{RA1}$$

Hence (from Equation 25 and Inequality RA1 after canceling out $\mathrm{Pr}_n(E_n)$):

$$-a\mathrm{Pr}(E_n|H_1) - \ldots - b\mathrm{Pr}(E_n|H_{n-1}) + c\mathrm{Pr}(E_n|H_{n+1}) + \ldots + d\mathrm{Pr}(E_n|H_m) \leq 0 \tag{26}$$

We know that hypotheses are jointly exhaustive, hence the changes in their probability sum to 0. That is: $-a - \ldots - b + c + \ldots + d = 0$. We then substitute c with $a + \ldots + b - \ldots - d$. After some rearrangements of Inequality 26, we obtain:

$$\begin{aligned} a\big(\mathrm{Pr}(E_n|H_{n+1}) - \mathrm{Pr}(E_n|H_1)\big) + \ldots + b\big(\mathrm{Pr}(E_n|H_{n+1}) - \mathrm{Pr}(E_n|H_{n-1})\big) \\ - \ldots - d\big(\mathrm{Pr}(E_n|H_{n+1}) - \mathrm{Pr}(E_n|H_m)\big) \leq 0 \end{aligned} \tag{27}$$

All terms are positive because $a, \ldots, b, \ldots, d$ are positive and $\mathrm{Pr}(E_n|H_{n+1}) > \mathrm{Pr}(E_n|H_i)$ for all $i, i < n+1$, and $\mathrm{Pr}(E_n|H_{n+1}) < \mathrm{Pr}(E_n|H_i)$ for all $i, i > n+1$. Inequality 27, and hence our reductio assumption (RA1), must therefore be false. We conclude our reductio with $\mathrm{Pr}_n(E_n) < \mathrm{Pr}_{n+1}(E_{n+1})$. We have now shown the first part of the Inequality 22.

Note that we can also show the first part of Inequality 23, i.e., $\mathrm{Pr}_n(E_n) > \mathrm{Pr}_{n+1}(E_{n+1})$ given $\mathrm{Pr}^*_n(E_n) < \mathrm{Pr}_n(E_n)$, by following the same method with different assumptions (proof omitted).

To prove that Inequality 22 holds, we need to also show that, given $\mathrm{Pr}_n(E_n) < \mathrm{Pr}^*_n(E_n)$, the following inequality holds:

$$\mathrm{Pr}_{n+1}(E_{n+1}) \leq \mathrm{Pr}^*_n(E_n) \tag{28}$$

To show Inequality 28, first note that because for all i, $\mathrm{Pr}_{n+1}(H_i) = \mathrm{Pr}^*_n(H_i)$, $\mathrm{Pr}_{n+1}(E_{n+1})$ may be calculated in the following way by the law of total probability:

$$\Pr_{n+1}(E_{n+1}) = \sum_{i=1}^{m} \Pr^*_n(H_i)(\Pr(E_n|H_i) \tag{29}$$

where m is the number of hypotheses. Recall the Jeffrey conditionalization rule for binary partitions of evidence. For all i, $\Pr^*_n(H_i)$ updates to:

$$\Pr^*_n(H_i) = \frac{\Pr^*_n(E_n)}{\Pr_n(E_n)}\Pr_n(H_i)\Pr(E_n|H_i) + \frac{\Pr^*_n(\neg E_n)}{\Pr_n(\neg E_n)}\Pr_n(H_i)\Pr(\neg E_n|H_i) \tag{JC}$$

Hence:

$$\begin{aligned}\Pr_{n+1}(E_{n+1}) = &\left(\frac{\Pr^*_n(E_n)}{\Pr_n(E_n)}\sum_{i=1}^{m}\Pr_n(H_i)\big(\Pr(E_n|H_i)\big)^2\right) + \\ &\left(\frac{\Pr^*_n(\neg E_n)}{\Pr_n(\neg E_n)}\sum_{i=1}^{m}\Pr_n(H_i)\Pr(\neg E_n|H_i)\Pr(E_n|H_i)\right)\end{aligned} \tag{30}$$

Recall that, for all i, $\Pr(\neg E_n|H_i) = 1 - \Pr(E_n|H_i)$, and that $\Pr_n(E_n) = \sum_{i=1}^{m}\Pr_n(H_i)\Pr(E_n|H_i)$. After some algebra we then obtain:

$$\begin{aligned}\Pr_{n+1}(E_{n+1}) = &\left(\frac{\Pr^*_n(E_n)}{\Pr_n(E_n)}\sum_{i=1}^{m}\Pr_n(H_i)\big(\Pr(E_n|H_i)\big)^2\right) \\ &- \left(\frac{\Pr^*_n(\neg E_n)}{\Pr_n(\neg E_n)}\sum_{i=1}^{m}\Pr_n(H_i)\big(\Pr(E_n|H_i)\big)^2\right) + \frac{\Pr^*_n(\neg E_n)}{\Pr_n(\neg E_n)}\Pr_n(E_n)\end{aligned} \tag{31}$$

Assume the following inequality for reductio ad absurdum:

$$\Pr_{n+1}(E_{n+1}) > \Pr^*_n(E_n) \tag{RA2}$$

After multiplying both sides of Inequality (RA2) by $\Pr_n(E_n)\Pr_n(\neg E_n)$, replacing all $\Pr^*_n(\neg E_n)$ with $1 - \Pr^*_n(E_n)$ and $\Pr_n(\neg E_n)$ with $1 - \Pr_n(E_n)$, and a few lines of algebra, we then obtain:

$$\Big(\Pr_n(E_n) - \Pr^*_n(E_n)\Big)\left(\Pr_n(E_n) - \sum_{i=1}^{m}\Pr_n(H_i)\big(\Pr(E_n|H_i)\big)^2\right) > 0 \tag{32}$$

We know that the left term is negative (recall our initial assumption $\Pr^*_n(E_n) > \Pr_n(E_n)$). We can also show that the second term is non-negative:

$$\mathrm{Pr}_n(E_n) - \sum_{i=1}^{m} \mathrm{Pr}_n(H_i)\big(\mathrm{Pr}(E_n|H_i)\big)^2 = \sum_{i=1}^{m} \mathrm{Pr}_n(H_i)\mathrm{Pr}(E_n|H_i) - \mathrm{Pr}_n(H_i)\big((\mathrm{Pr}(E_n|H_i)\big)^2 = \sum_{i=1}^{m} \mathrm{Pr}_n(H_i)\mathrm{Pr}(E_n|H_i)\mathrm{Pr}(\neg E_n|H_i) \geq 0 \quad (33)$$

Hence, Inequality 32 cannot be true. We thus conclude that our reductio assumption (RA2) is false, so $\mathrm{Pr}_{n+1}(E_{n+1}) \leq \mathrm{Pr}^*_n(E_n)$ is true. Inequality 22, therefore, holds.

Note that we can also prove $\mathrm{Pr}_{n+1}(E_{n+1}) > \mathrm{Pr}^*_n(E_n)$ given $\mathrm{Pr}_n(E_n) \geq \mathrm{Pr}^*_n(E_n)$ by following the same method (proof omitted). Inequality 23, therefore, also holds. This concludes our proof of Theorem 4. □