

# Awareness Growth in Bayesian Networks

July 28, 2022

*Wordcount (including footnotes): 5,392*

We examine different counterexamples to Reverse Bayesianism, a popular theory that addresses the problem of awareness growth. We agree with the general skepticism toward Reverse Bayesianism, but submit that the problem of awareness growth cannot be tackled in an algorithmic manner, because subject-matter, structural assumptions need to be made explicit. Thanks to their ability to express probabilistic dependencies, we illustrate how Bayesian networks can help to model awareness growth in the Bayesian framework.

## 1 Introduction

Learning is modeled in the Bayesian framework by the rule of conditionalization. This rule posits that the agent's new degree of belief in a proposition  $H$  after a learning experience  $E$  should be the same as the agent's old degree of belief in  $H$  conditional on  $E$ . That is,

$$P^E(H) = P(H|E),$$

where  $P()$  represents the agent's old degree of belief (before the learning experience  $E$ ) and  $P^E()$  represents the agent's new degree of belief (after the learning experience  $E$ ).

Both  $E$  and  $H$  belong to the agent's algebra of propositions. This algebra models the agent's awareness state, the propositions taken to be live possibilities. Conditionalization never modifies the algebra and thus makes it impossible for an agent to learn something they have never thought about. Even before learning about  $E$ , the agent must already have assigned a degree of belief to any proposition conditional on  $E$ . This picture commits the agent to the specification of their 'total possible future experience' (Howson, 1976), as though learning was confined to an 'initial prison' (Lakatos, 1968).

1 But, arguably, the learning process is more complex than what conditionalization allows.  
2 Not only do we learn that some propositions that we were entertaining are true or false, but  
3 we may also learn new propositions that we did not entertain before. Or we may entertain  
4 new propositions—without necessarily learning that they are true or false—and this change  
5 in awareness may in turn change what we already believe. How should this more complex  
6 learning process be modeled by Bayesianism? Call this the problem of awareness growth.

7 The algebra of propositions need not be so narrowly construed that it only contains proposi-  
8 tions that are presently under consideration. The algebra may also contain propositions which,  
9 though outside the agent’s present consideration, are still the object, perhaps implicitly, of  
10 certain dispositions to believe.<sup>1</sup> But even this expanded algebra will have to be revised sooner  
11 or later. The algebra of propositions could in principle contain anything that could possibly be  
12 conceived, expressed, thought of. Such a rich algebra would not need to change at any point,  
13 but this is an implausible model of ordinary agents with bounded resources such as ourselves.

14 Critics of Bayesianism and sympathizers alike have been discussing the problem of awareness  
15 growth under different names for quite some time, at least since the eighties. This problem arises  
16 in a number of different contexts, for example, new scientific theories (Chihara, 1987; Earman,  
17 1992; Glymour, 1980), language changes and paradigm shifts (Williamson, 2003), and theories  
18 of induction (Zabell, 1992). A proposal that has attracted considerable scholarly attention  
19 in recent years is Reverse Bayesianism (Bradley, 2017; Karni & Vierø, 2015; Wenmackers  
20 & Romeijn, 2016). The idea is to model awareness growth as a change in the algebra while  
21 ensuring that the proportions of probabilities of the propositions shared between the old and  
22 new algebra remain the same in a sense to be specified.

23 Let  $\mathcal{F}$  be the initial algebra of propositions and let  $\mathcal{F}^+$  the algebra after the agent’s awareness  
24 state has grown. Both algebras contain the contradictory and tautologous propositions  $\perp$  and  
25  $\top$ , and they are closed under connectives such as disjunction  $\vee$ , conjunction  $\wedge$  and negation  $\neg$ .  
26 Denote by  $X$  and  $X^+$  the subsets of these algebras that contain only basic propositions, namely  
27 those without connectives. **Reverse Bayesianism** posits that the ratio of probabilities for any  
28 basic propositions  $A$  and  $B$  in both  $X$  and  $X^+$ —the basic propositions shared by the old and

---

<sup>1</sup>Roussos (2021) notes that, for the sake of clarity, the problem of awareness growth should only address propositions which agents are *truly* unaware of (say new scientific theories), not propositions that were temporarily forgotten or set aside. This is a helpful clarification to keep in mind, although the recent literature on the topic does not make a sharp distinction between true unawareness and temporary unawareness.

1 new algebra—remain constant through the process of awareness growth:

$$\frac{P(A)}{P(B)} = \frac{P^+(A)}{P^+(B)},$$

2 where  $P()$  represents the agent's degree of belief before awareness growth and  $P^+()$  represents  
3 the agent's degree of belief after awareness growth.

4 Reverse Bayesianism is an elegant theory that manages to cope with a seemingly intractable  
5 problem. As the awareness state of an agent grows, the agent would prefer not to throw  
6 away completely the epistemic work they have done previously. The agent may desire to retain  
7 as much of their old degrees of beliefs as possible. Reverse Bayesianism provides a simple  
8 recipe to do that. It also coheres with the conservative spirit of Bayesian conditionalization  
9 which preserves the old probability distribution conditional on what is learned.

10 Unfortunately, Reverse Bayesianism does not deliver the intuitive results in all cases. There  
11 is no shortage of counterexamples against it in the recent philosophical literature (Mathani,  
12 2020; Steele & Stefánsson, 2021). In addition, attempts to extend traditional arguments in  
13 defense of Bayesian conditionalization to the case of awareness growth seem to hold little  
14 promise (Pettigrew, forthcoming). If the consensus in the literature is that Reverse Bayesianism  
15 is not the right theory of awareness growth, what theory (if any) should replace it?

16 Here we offer a diagnosis of what is wrong with Reverse Bayesianism and outline an  
17 alternative proposal. The problem of awareness growth—we hold—cannot be tackled in an  
18 algorithmic manner because subject-matter assumptions, both probabilistic and structural, need  
19 to be made explicit. So any theory of awareness growth cannot be a purely formal theory. This  
20 does not mean, however, we should give up on probability theory altogether. Thanks to its  
21 ability to express probabilistic dependencies, we think that the theory of Bayesian networks  
22 can help to model awareness growth in the Bayesian framework. We illustrate this claim as we  
23 examine different counterexamples to Reverse Bayesianism.

## 24 **2 Counterexamples to Reverse Bayesianism**

25 A common set of cases of awareness growth are usually referred to in the literature by the label  
26 *awareness expansion*. A precise definition of expansion can be tricky to provide, but a rough  
27 characterization will suffice here. Suppose, as is customary, propositions are interpreted as sets

1 of possible worlds, where the set of all possible worlds is the possibility space. Awareness  
2 expansion occurs when a new proposition is added to the algebra and its interpretation includes  
3 possible worlds not in the original possibility space. So the addition of the new proposition  
4 causes the possibility space to expand.

5 Perhaps the most straightforward example of awareness expansion occurs when you become  
6 aware of a new explanation for the evidence at your disposal which you had not considered  
7 before. This can happen in many fields of inquiry: medicine, law, science, everyday affairs.  
8 Here is a scenario by Steele & Stefánsson (2021):

9 FRIENDS: Suppose you happen to see your partner enter your best friend's house  
10 on an evening when your partner had told you she would have to work late. At  
11 that point, you become convinced that your partner and best friend are having  
12 an affair, as opposed to their being warm friends or mere acquaintances. You  
13 discuss your suspicion with another friend of yours, who points out that perhaps  
14 they were meeting to plan a surprise party to celebrate your upcoming birthday—a  
15 possibility that you had not even entertained. Becoming aware of this possible  
16 explanation for your partner's behaviour makes you doubt that she is having an  
17 affair with your friend, relative, for instance, to their being warm friends. (Steele  
18 & Stefánsson, 2021, sec. 5, Example 2)

19 Initially, the algebra contained the hypotheses 'my partner and my best friend met to have  
20 an affair' (*Affair*) and 'my partner and my best friend met as friends or acquaintances'  
21 (*Friends/acquaintances*). These were the only explanations you considered for your evidence,  
22 namely that your partner and your best friend met one night without telling you (*Secretive*).  
23 Given this evidence, *Affair* is more probable than *Friends/acquaintances*:<sup>2</sup>

$$P(Affair|Secretive) > P(Friends/acquaintances|Secretive). \quad (>)$$

24 When the algebra changes, a new hypothesis is added which you had not considered before:  
25 your partner and your best friends met to plan a surprise party for your upcoming birthday  
26 (*Surprise*). This hypothesis seems to be a much better explanation. So, given the same evidence,

---

<sup>2</sup>This assumes that the prior probabilities of the two hypotheses were not strongly skewed in one direction. If you were initially nearly certain your partner could not possibly have an affair, even the fact they behaved very secretly or lied to you might not affect the probability of the two hypotheses.

1 the hypothesis *Surprise* is more likely than the hypothesis *Affair*:

$$P^+(Surprise|Secretive) > P^+(Affair|Secretive).$$

2 And since *Surprise* implies *Friends/acquaintances*—after all, in order to prepare a surprise  
3 party, your partner and best friend have to be at least acquaintances— *Friends/acquaintances* is  
4 now more likely than *Affair*:

$$P^+(Affair|Secretive) < P^+(Friends/acquaintances|Secretive). \quad (<)$$

5 As Steele & Stefánsson note, the conjunction of ( $>$ ) and ( $<$ ) violates Reverse Bayesianism.  
6 But, they also note, a quick fix is available. It is reasonable to suppose that no change in the  
7 probability assignments should occur so long as we confine ourselves to the old probability  
8 space. With this in mind, consider the following condition, called **Awareness Rigidity**:

$$P^+(A|T^*) = P(A),$$

9 where  $T^*$  corresponds to a proposition that picks out, from the vantage point of the new  
10 awareness state, the entire possibility space *before* the episode of awareness growth. Awareness  
11 rigidity establishes that, once a suitable proposition  $T^*$  is identified, the old probability assign-  
12 ments remain unchanged conditional on  $T^*$ . In our running example,  $\neg Surprise$  is the suitable  
13 proposition  $T^*$ : *that* there was no surprise party in the making picks out the original  
14 possibility space. After all, before awareness growth, the eventuality that your partner and best  
15 friend could be organizing a surprise for you had been tacitly ruled out. So, conditional on  
16  $\neg Surprise$ , no probability assignment should change, including the probability of *Affair*. Thus,

$$P^+(Affair|Secretive \& \neg Surprise) > P^+(Friends/acquaintances|Secretive \& \neg Surprise).$$

17 Steele and Stefansson, however, go on to show that Awareness Rigidity does not hold in  
18 other cases, what they call *awareness refinement*. For consider this scenario:

19 MOVIES: Suppose you are deciding whether to see a movie at your local cinema.  
20 You know that the movie's predominant language and genre will affect your  
21 viewing experience. The possible languages you consider are French and German

1 and the genres you consider are thriller and comedy. But then you realise that,  
2 due to your poor French and German skills, your enjoyment of the movie will  
3 also depend on the level of difficulty of the language. Since it occurs to you that  
4 the owner of the cinema is quite simple-minded, you are, after this realisation,  
5 much more confident that the movie will have low-level language than high-level  
6 language. Moreover, since you associate low-level language with thrillers, this  
7 makes you more confident than you were before that the movie on offer is a thriller  
8 as opposed to a comedy. (Steele & Stefánsson, 2021, sec. 5, Example 3)

9 This is a case of refinement. For the new proposition added to the algebra induces a more fine-  
10 grained partition of the possibility space. You initially categorized movies by just language and  
11 genre, and then you refined your categorization by adding another variable, level of difficulty.  
12 Without considering language difficulty, you assigned the same probability to the hypotheses  
13 *Thriller* and *Comedy*. But learning that the owner was simple-minded made you think that  
14 the level of linguistic difficulty must be low and the movie most likely a thriller rather than a  
15 comedy (perhaps because thrillers are simpler—linguistically—than comedies). So, against  
16 Reverse Bayesianism, MOVIES violates the condition  $\frac{P(Thriller)}{P(Comedy)} = \frac{P^+(Thriller)}{P^+(Comedy)}$ .

17 The counterexample also violates Awareness Rigidity. For consider a proposition that picks  
18 out the entire possibility space, for example,  $Thriller \vee Comedy$ .<sup>3</sup> Awareness Rigidity would  
19 require that  $P(Thriller) = P^+(Thriller|Thriller \vee Comedy)$ . But MOVIES does not satisfy this  
20 equality since the probability of *Thriller* has gone up.<sup>4</sup>

21 We believe that what is wrong with principles such as Reverse Bayesianism or Awareness  
22 Rigidity is that they are purely formal. Any principle intended to govern awareness growth  
23 should be grounded in the subject-matter information underlying the scenario at hand. This  
24 subject-matter takes many forms. In FRIENDS, awareness expansion does not change the basic  
25 presupposition that someone's behavior must have a reason. In MOVIES, awareness refinement  
26 does not change the fact that certain characteristics, such as language, difficulty or genre, may  
27 influence one's decision to select a movie for showing rather than another. So, in modeling

<sup>3</sup>Since MOVIES is a case of refinement,  $Thriller \vee Comedy$  picks out the entire possibility space both before and after awareness growth.

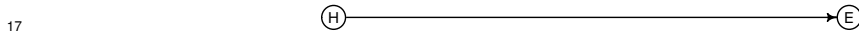
<sup>4</sup>MOVIES can be split into two episodes. In the first, you entertain a new variable besides language and genre, namely the language difficulty of the movie. In the second episode, you learn something you did not consider before, namely that the owner is simple-minded. Could Reverse Bayesianism still work for the first episode, but not the second? Steele and Stefánsson do not address this question explicitly, but insist that no matter the answer both episodes are instances of awareness growth. We agree with them on this point. Awareness growth is both *entertaining* a new proposition not in the initial awareness state of the agent and *learning* a new proposition.

1 cases of awareness growth, a formalism that can also represent this subject-information is  
 2 needed. In what follows, we illustrate how Bayesian networks can serve this purpose. They  
 3 may not be the perfect formalism, but they certainly are a promising starting point.

### 4 3 Expansion with Bayesian Networks

5 A Bayesian network is a compact formalism to represent probabilistic dependencies. It consists  
 6 of a direct acyclic graph (DAG) accompanied by a probability distribution. The nodes in  
 7 the graph represent random variables that can take different values. We will use ‘nodes’ and  
 8 ‘variables’ interchangeably. The nodes are connected by arrows, but no loops are allowed,  
 9 hence the name direct acyclic graph. Bayesian networks are relied upon in many fields, but have  
 10 been rarely deployed to model awareness growth (the exception is Williamson (2003)). We  
 11 think instead they are a good framework for this purpose. Awareness growth can be modeled as  
 12 a change in the graphical network—nodes and arrows are changed, added or erased—as well  
 13 as a change in the probability distribution from the old to the new network. In this section, we  
 14 focus on cases of awareness expansion and turn to refinement in the next section.

15 Recall the scenario FRIENDS from before. To model it using a Bayesian network, we start  
 16 with this graph, which is the usual hypothesis-evidence idiom (Fenton, Neil, & Lagnado, 2013):



18 where  $H$  is the hypothesis node and  $E$  the evidence node. If an arrow goes from  $H$  to  $E$ , the  
 19 full probability distribution associated with the Bayesian network is defined by two probability  
 20 tables.<sup>5</sup> One table defines the prior probabilities  $P(H = h)$  for all the states (or values) of  $H$ ,  
 21 and another table defines the conditional probabilities of the form  $P(E = e|H = h)$ , where  
 22 uppercase letters represent the variables (nodes) and lower case letters represent the states  
 23 (or values) of these variables. These two probability tables are sufficient to specify the full  
 24 probability distribution. The other probabilities—say  $P(E = e)$ ,  $P(H = h|E = e)$ , etc.—follow  
 25 by simply applying the probability axioms.

26 Initially, before awareness growth, the hypothesis node  $H$  takes only two states,  
 27 *friends/acquaintances* and *affair*. These two states are meant to be exhaustive, so *affair*

<sup>5</sup>A major point of contention in the interpretation of Bayesian networks is is the meaning of the directed arrows. They could be interpreted causally—as though the direction of causality proceeds from the events described by the hypothesis to event described by the evidence—but they need not be; see footnote 13.

1 functions as the negation of *friends/acquaintances*, and vice versa. After awareness growth—in  
 2 fact, awareness expansion—the two states are no longer exhaustive. Node  $H$  now has a third  
 3 state: *surprise*. So, expansion simply consists in the addition of an extra state in one of the  
 4 nodes of the network. The rest of the structure of the network remains intact.

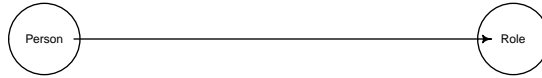
5 In FRIENDS, it is plausible to assume that the extra state added to the upstream node  $H$  does  
 6 not change the plausibility to the two hypotheses considered at first relative to one another.  
 7 Even after awareness expansion, the evidence *secretive* makes better sense in light of the  
 8 hypothesis *affair* compared to *friends/acquaintances*, despite the fact that the novel hypothesis  
 9 *surprise* makes better sense of the evidence overall. Thus, for all values  $e$  and  $h$  of upstream  
 10 node  $H$  and downstream node  $E$  in the old network, the following constraint holds:

$$\frac{P(E = e|H = h)}{P(E = e|H = h)} = \frac{P^+(E = e|H = h)}{P^+(E = e|H = h)}. \quad (C)$$

11 Constraint (C) is a restricted variant of Reverse Bayesianism that only applies to the conditional  
 12 probabilities in the probability table for the evidence-hypothesis Bayesian network.

13 In TENANT, it is natural to assign 1/3 to *Landlord*, *Bob* and *Other* after awareness growth.  
 14 That someone is singing in the shower is evidence that someone must be in there, but without  
 15 any more discriminating evidence, each person should be assigned the same probability.  
 16 Consequently, a probability of 2/3 should be assigned to *Tenant*. On this picture, the proportion  
 17 of *Landlord* to *Tenant* changes from 1:1 (before awareness growth) to 1:2 (after awareness  
 18 growth).

19 Bayesian networks can help to model this scenario. We start with the following graph:



21 Initially, the upstream node *Person* has two possible states, representing who is in the bathroom  
 22 singing: *landlord-person* and *bob*. To simplify things, the assumption here is that the evidence  
 23 of singing has already ruled out the possibility that no one would be in the shower. The  
 24 downstream node *Role* has also two values, *landlord* and *tenant*. After your awareness grows,  
 25 the upstream node *Person* should now have one more possible state, *other*.

26 The scenarios FRIENDS and TENANT are structurally identical as far as their modeling using  
 27 Bayesian networks. So, if constraint (C) holds in FRIENDS, as seen earlier, it must also hold in



1 TENANT. This is precisely what happens. For these ratios remain fixed:

$$\frac{P(\text{Role} = \text{landlord} | \text{Person} = \text{landlord})}{P(\text{Role} = \text{landlord} | \text{Person} = \text{bob})} = \frac{P^+(\text{Role} = \text{landlord} | \text{Person} = \text{landlord})}{P^+(\text{Role} = \text{landlord} | \text{Person} = \text{bob})} = 1/0$$

$$\frac{P(\text{Role} = \text{tenant} | \text{Person} = \text{landlord})}{P(\text{Role} = \text{tenant} | \text{Person} = \text{bob})} = \frac{P^+(\text{Role} = \text{tenant} | \text{Person} = \text{landlord})}{P^+(\text{Role} = \text{tenant} | \text{Person} = \text{bob})} = 0/1$$

2 The upshot of this discussion is this. Awareness expansion can be modeled by changes in the  
3 Bayesian network used to represent the epistemic state of the agent. The structure of the network  
4 does not change, but values are added to the upstream nodes. (We will consider changes to the  
5 downstream node shortly.) This addition can be carried while satisfying a restricted version of  
6 Reverse Bayesian, what we called constraint (C). This constraint outperforms both Reverse  
7 Bayesianism (which fails in FRIENDS) and Awareness Rigidity (which fails in TENANT).

8 What would happen is the extra value is added to the downstream node in the hypothesis-  
9 evidence network? Consider a variation of FRIENDS. Suppose that the evidence node  $E$  could  
10 initially take only two values, say *Secretive* and *Public*. You then realize the evidence node  
11 could also take a third value, say *Ambiguous*. This realization mandate a change in the old  
12 conditional probabilities  $P(E = e | H = h)$ . Since *Secretive* and *Public* were initially considered  
13 exhaustive,

14 Since the change is downstream the old conditional probabilities  $P(E = e | H = h)$  would  
15 remain the same, and thus *a fortiori*, constraint (C) would be satisfied.

16 outperformes the same result as Awareness Rigidity, but more easily applies to the conditional  
17 probabilities that are needed in the probability tables of Bayesian Networks. The challenge  
18 now is to develop a systematic method to determine when constraint (C) is satisfied and when  
19 it fails. The structure of the Bayesian network will be our guide. This will afford us a firmer  
20 foundation to develop a general theory of awareness growth.

21 But before turning to cases of refinement, it is instructive to see that Awareness Rigidity  
22 does not even work for all cases of expansion. For consider the following scenario by Mathani  
23 (2020):

24 TENANT: You are staying at Bob's flat which he shares with his landlord. You

1 know that Bob is a tenant, and that there is only one landlord, and that this landlord  
2 also lives in the flat. In the morning you hear singing coming from the shower  
3 room, and you try to work out from the sounds who the singer could be. At this  
4 point you have two relevant propositions that you consider possible ... *Landlord*  
5 standing for the possibility that the landlord is the singer, and *Bob* standing for  
6 the possibility that Bob is the singer ... Because you know that Bob is a tenant  
7 in the flat, you also have a credence in the proposition *Tenant* that the singer is a  
8 tenant. Your credence in *Tenant* is the same as your credence in *Bob*, for given  
9 your state of awareness these two propositions are equivalent ... Now let's suppose  
10 the possibility suddenly occurs to you that there might be another tenant living in  
11 the same flat (*Other*).

12 Initially, you thought the singer could either be the landlord or Bob, the tenant. Then you  
13 come to the realization that a third person could be the singer, another tenant. This scenario  
14 is a bit more complicated than FRIENDS. For one thing, the possibility that there could be  
15 a third person in the shower—besides Bob or the landlord—is a novel explanation for why  
16 you hear singing in the shower. So TENANT seems to be a standard case of expansion. The  
17 expansion in awareness, however, goes along with an interesting conceptual shift. Before  
18 awareness expansion, that Bob is in the shower and that a tenant is in the shower are equivalent  
19 descriptions, but after the expansion, this equivalence breaks down.

20 This scenario is problematic for both Reverse Bayesianism and Awareness Rigidity. Suppose,  
21 after you hear singing in the shower, you become sure someone is in there, but you cannot  
22 tell who. So  $P(\text{Landlord}) = P(\text{Bob}) = 1/2$ , and since *Bob* and *Tenant* are equivalent, also  
23  $P(\text{Tenant}) = 1/2$ . Now, *Landlord*, *Bob* and *Tenant* are all propositions that you were originally  
24 aware of, and thus Reverse Bayesianism requires that their probabilities should remain in the  
25 same proportion after your awareness grows. But note that *Other* entails *Tenant* and *Bob* and  
26 *Other* are disjoint, so it follows that  $P^+(\text{Other})$  must have zero probability. If  $P^+(\text{Other})$  were  
27 greater than zero, the proportion of the probability of *Tenant* to *Landlord* (or the proportion  
28 of the probability of *Bob* to *Landlord*) should change. Awareness Rigidity is no of help  
29 either because it would require that  $P^+(\text{Landlord}|\text{Landlord} \vee \text{Tenant}) = P^+(\text{Bob}|\text{Landlord} \vee$   
30  $\text{Tenant})$  both equal  $1/2$ , thus forcing  $P^+(\text{Other}|\text{Landlord} \vee \text{Tenant})$  to zero.

## 1 4 Refinement

2 We turn now from cases of awareness expansion to cases of awareness refinement. In the  
3 framework of Bayesian networks, expansion consisted in added values to nodes in the network.  
4 Refinement, instead, can be modeled by adding nodes to the network without necessarily add  
5 any new values to the existing nodes. Intuitively, refinement takes place when an epistemic  
6 agent acquires a more-fined grained picture of the situation, say instead of thinking that the  
7 political spectrum is divided into liberal and conservatives, the political spectrum can be further  
8 divided into traditional-liberal, new-liberal, traditional-conservative and new-conservative. The  
9 political spectrum is still divided into liberal and conservative—not expansion occurred—but  
10 these two categories have been further refined.

11 Although there is no shortage of counterexamples to Reverse Bayesianism when it comes  
12 to awareness refinement, we begin with our own. This will allow us to underscore the role  
13 of subject-matter assumptions in theorizing about awareness growth. Consider the following  
14 scenario:

15 LIGHTING: You have evidence that favors a certain hypothesis, say a witness saw  
16 the defendant around the crime scene. You give some weight to this evidence.  
17 In your assessment, that the defendant was seen around the crime scene (your  
18 evidence) raises the probability that the defendant was actually there (your hypoth-  
19 esis). But now you ask, what if it was dark when the witness saw the defendant?  
20 In light of your realization that it could have been dark, you wonder whether (and  
21 if so how) you should change the probability that you assigned to the hypothesis  
22 that the defendant was around the crime scene.

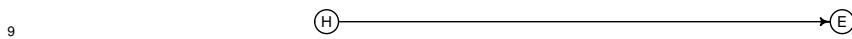
23 As your awareness grows, you do not learn anything specific about the lighting conditions,  
24 neither that they were bad nor that they were good. You simply wonder what they were, a  
25 variable you had previously not considered. So no Bayesian updating takes place in the strict  
26 sense, although broadly speaking some new information has been introduced.<sup>6</sup> Something has

---

<sup>6</sup>HERE EXPLAIN DIFFERENCE WITH STEELE AND STEFANSSON. The process of awareness growth in LIGHTING adds only one extra variable, lighting conditions, while MOVIES adds two extra variables, language difficulty and whether the owner is simple-minded or not. Further, MOVIES contains a clear-cut case of learning, that the owner *is* simple-minded. This is not so in LIGHTING. Strictly speaking, you are learning that it is *possible* that the lighting conditions were bad. However, you are not conditioning on the proposition ‘the lighting conditions were bad’ or ‘the lighting conditions were good’. So you are not learning about the lighting conditions in the sense in which learning is understood in this paper.

1 changed in your epistemic state—you have a more fine-grained assessment of what could have  
 2 happened—but it is not clear what you should do in this scenario. Since the lighting conditions  
 3 could have been bad but could also have been good, perhaps you should just stay put until you  
 4 learn something more specific.

5 In what follows, we illustrate how Bayesian networks helps to model what is going on in  
 6 LIGHTING and conclude that you should probably revise downward your confidence in the  
 7 hypothesis that the defendant was around the crime scene. The starting point of our analysis is  
 8 the usual hypothesis-evidence idiom, repeated below for convenience:

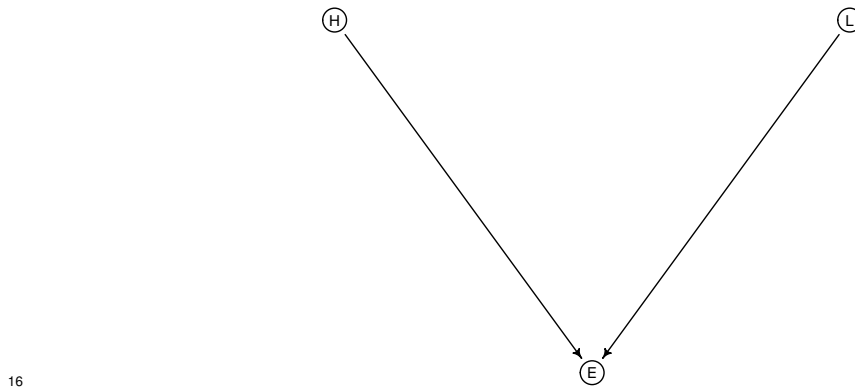


10 Since you trust the evidence, you think that the evidence is more likely under the hypothesis  
 11 that the defendant was present at the crime scene than under the alternative hypothesis:

$$P(E=seen|H=present) > P(E=seen|H=absent)$$

12 The inequality is a qualitative ordering of how plausible the evidence is in light of competing  
 13 hypotheses. No matter the numbers, by the probability calculus, it follows that the evidence  
 14 raises the probability of the hypothesis  $H=present$ .

15 Now, as you wonder about the lighting conditions, the graph should be amended:



17 where the node  $L$  can have two values,  $L=good$  and  $L=bad$ . Commonsense as well as psycho-  
 18 logical findings suggest that when the visibility deteriorates, people's ability to identify faces  
 19 worsen. So a plausible way to modify your assessment of the evidence is as follows:

$$P^+(E=seen|H=present \wedge L=good) > P^+(E=seen|H=absent \wedge L=good)$$

$$P^+(E=seen|H=present \wedge bad) = P^+(E=seen|H=absent \wedge L=bad)$$

2 In words, if the lighting conditions were good, you still trust the evidence like you did before  
 3 (first line), but if the lighting conditions were bad, you regard the evidence as no better than  
 4 chance (second line). These probabilistic constraints are plausible, but should ultimately be  
 5 grounded on verifiable empirical regularities.

6 Despite the change in awareness, you have not learned anything in the strict sense. Your  
 7 new stock of evidence does not contain neither the information that the lighting conditions  
 8 were bad nor that they were good. But the Bayesian network structure that represents your  
 9 epistemic state is now more fine-grained. The network contains the new variable  $L$  which it  
 10 did not contain prior to the episode of awareness growth. In addition—and this is the crucial  
 11 point—the new variable bears certain *structural relationships* with the variables  $H$  and  $E$ . The  
 12 graphical network represents a direct probabilistic dependency between the lighting conditions  
 13  $L$  and the witness sensory experience  $E$ , but does not allow for any direct dependency between  
 14 the lighting conditions and the fact that the defendant was (or was not) at the crime scene.  
 15 There is no direct arrow between the nodes  $L$  and  $H$ . This structure of dependencies captures  
 16 our causal intuitions about the scenario: the lighting conditions do affect what the witness  
 17 could see, but do not directly affect what the defendant might have have done.

18 Without Bayesian networks, episodes of awareness growth could only be modeled by the  
 19 addition of new propositions that were not previously in the algebra. But this approach would  
 20 fail to capture crucial information. When awareness growth takes place against the background  
 21 of an intuitive causal structure of the world—as in the case of LIGHTING—this structure should  
 22 also be modeled. Bayesian networks offer a formal framework that can do precisely that.

23 This model of causal structure can now guide us to decide whether the restricted version of  
 24 Reverse Bayesianism, what we called constraint (C), holds in this scenario. Specifically, we  
 25 need to assess whether the following holds:

$$\frac{P(E = seen|H = present)}{P(E = seen|H = absent)} = \frac{P^+(E = seen|H = present)}{P^+(E = seen|H = absent)}.$$

26 The question here is whether you should assess the evidence at your disposal—that the witness  
 27 saw the defendant at the crime scene—any differently than before. As noted earlier, without a

1 clear model of the scenario, it might seem that you should simply stay put. After all, besides  
 2 the sensory experience of the witness, you have gained no novel information about the lighting  
 3 conditions. Should you thus conclude that the evidence has the same value before and after the  
 4 realization that lighting could have been bad?

5 The evidence would have the same value if the likelihood ratios associated with it relative to  
 6 the competing hypotheses were the same before and after awareness growth. But, in changing  
 7 the probability function from  $P()$  to  $P^+$ , it would be quite a coincidence if this were true. In  
 8 our example, many possible probability assignments violate this equality. If before awareness  
 9 growth you thought the evidence favored the hypothesis  $H=present$  to some extent, after the  
 10 growth in awareness, the evidence is likely to appear less strong.<sup>7</sup> If this is correct, this outcome  
 11 violates constraint (C). Reverse Bayesianism is also violated since the ratio of the probabilities

---

<sup>7</sup>By the law of total probability, the right hand side of the equality in (C) should be expanded, as follows:

$$\frac{P^+(E=e|H=h)}{P^+(E=e|H=h')} = \frac{P^+(E=seen \wedge L=good|H=present) + P^+(E=seen \wedge L=bad|H=present)}{P^+(E=seen \wedge L=good|H=absent) + P^+(E=seen \wedge L=bad|H=absent)}.$$

For concreteness, let's use some numbers:

$$P(E=seen|H=present) = P^+(E=seen|H=present \wedge L=good) = .8$$

$$P(E=seen|H=absent) = P^+(E=seen|H=absent \wedge L=good) = .4$$

$$P^+(E=seen|H=present \wedge L=bad) = P^+(E=seen|H=absent \wedge L=bad) = .5.$$

$$P^+(L=bad) = P^+(L=good) = .5.$$

So the ratio  $\frac{P(E=seen|H=present)}{P(E=seen|H=absent)}$  equals 2. After the growth in awareness, the ratio  $\frac{P^+(E=seen|H=present)}{P^+(E=seen|H=absent)}$  will drop to  $\frac{.65}{.45} \approx 1.44$ . The calculations here rely on the dependency structure encoded in the Bayesian network (see starred step below).

$$\begin{aligned} P^+(E=seen|H=present) &= P^+(E=seen \wedge L=good|H=present) + P^+(E=seen \wedge L=bad|H=present) \\ &= P^+(E=seen|H=present \wedge L=good) \times P^+(L=good|H=present) \\ &\quad + P^+(E=seen|H=present \wedge L=bad) \times P^+(L=bad|H=present) \\ &= * P^+(E=seen|H=present \wedge L=good) \times P^+(L=good) \\ &\quad + P^+(E=seen|H=present \wedge L=bad) \times P^+(L=bad) \\ &= .8 \times .5 + .5 \times .5 = .65 \end{aligned}$$

$$\begin{aligned} P^+(E=seen|H=absent) &= P^+(E=seen \wedge L=good|H=absent) + P^+(E=seen \wedge L=bad|H=absent) \\ &= P^+(E=seen|H=absent \wedge L=good) \times P^+(L=good|H=absent) \\ &\quad + P^+(E=seen|H=absent \wedge L=bad) \times P^+(L=bad|H=absent) \\ &= * P^+(E=seen|H=absent \wedge L=good) \times P^+(L=good) \\ &\quad + P^+(E=seen|H=absent \wedge L=bad) \times P^+(L=bad) \\ &= .4 \times .5 + .5 \times .5 = .45 \end{aligned}$$

This argument can be repeated with many other numerical assignments.

1 of  $H=present$  to  $E=seen$ , before and after awareness growth, has changed:

$$\frac{P^{E=seen}(H=present)}{P^{E=seen}(E=seen)} \neq \frac{P^{+,E=seen}(H=present)}{P^{+,E=seen}(E=seen)},$$

2 where  $P^{E=seen}()$  and  $P^{+,E=seen}()$  represent the agent's degrees of belief, before and after aware-  
3 ness growth, updated by the evidence  $E=seen$ .<sup>8</sup>

4 The general lesson to be learned here has to do with the importance of formalizing structural  
5 assumptions and the role of Bayesian networks in modeling awareness growth. Modeling those  
6 structural assumptions allows us to see that constraint (C)—as well as Reverse Bayesianism  
7 more generally—fails here. To strengthen this point, consider this variation of the LIGHTING  
8 scenario:

9 VERACITY: A witness saw that the defendant was around the crime scene and  
10 you initially took this to be evidence that the defendant was actually there. But  
11 then you worry that the witness might be lying or misremembering what happened.  
12 Perhaps, the witness was never there, made things up or mixed things up. Should  
13 you reassess the evidence at your disposal? If so, how?

14 It might seem that this scenario is no different from LIGHTING. The realization that lighting  
15 could be bad should make you less confident in the truthfulness of the sensory evidence. And  
16 the same conclusion should presumably follow from the realization that the witness could be  
17 lying. So both scenarios would be counterexamples to Reverse Bayesianism. But, upon closer  
18 scrutiny, things are not that simple. To run the two scenarios together would be a mistake.

19 The evidence at your disposal in LIGHTING is the sensory evidence—the experience of  
20 seeing—and the possibility of bad lighting does affect the quality of your visual experience.  
21 So, if lighting was indeed bad, this would warrant lowering your confidence in the truthfulness  
22 of the visual experience. But the possibility of lying in VERACITY does not affect the quality  
23 of the visual experience in and of itself, although it affects the quality of the *reporting* of  
24 that experience. So, if the witness did lie, this would not warrant lowering your confidence

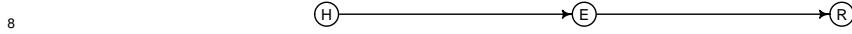
---

<sup>8</sup>The scenario also violates Awareness Rigidity which requires that  $P^+(A|T^*) = P(A)$ , where  $T^*$  corresponds to a proposition that picks out, from the vantage point of the new awareness state, the entire possibility space before the episode of awareness growth. In LIGHTING, however,  $T^*$  does not change, so Awareness Rigidity would require that  $P^+(A) = P(A)$ , and instead in the scenario, we have

$$P^+(H=present|E=seen) \neq P(H=present|E=seen).$$

1 in the truthfulness of the visual experience, only in the truthfulness of the reporting of that  
 2 experience. The distinction between the visual experience and its reporting is crucial here.  
 3 Bayesian networks help to model this distinction precisely, and then see why LIGHTING and  
 4 VERACITY are structurally different.

5 The graphical network should initially look like the initial DAG for LIGHTING, consisting  
 6 of the hypothesis node  $H$  upstream and the evidence node  $E$  downstream. As your awareness  
 7 grows, the graphical network should be updated by adding another node  $R$  further downstream:



9 As before, the hypothesis node  $H$  bears on the whereabouts of the defendant and has two values,  
 10  $H=present$  and  $H=absent$ . Note the difference between  $E$  and  $R$ . The evidence node  $E$  bears  
 11 on the visual experience had by the witness. The reporting node  $R$ , instead, bears on what the  
 12 witness reports to have seen. The chain of transmission from ‘visual experience’ to ‘reporting’  
 13 may fail for various reasons, such as lying or misremembering.

14 In VERACITY, the conditional probabilities,  $P(E = e | H = h)$  should be the same as  $P^+(E =$   
 15  $e | H = h)$  for any values  $e$  and  $h$  of the variables  $H$  and  $E$  that are shared before and after  
 16 awareness growth. In comparing the old and new Bayesian network, this equality falls out  
 17 from their structure, as the connection between  $H$  and  $E$  remains unchanged. Thus, constraint  
 18 (C)—along with Reverse Bayesianism—is perfectly fine in scenarios such as VERACITY.

19 This does not mean that the assessment of the probability of the hypothesis  $H=present$  should  
 20 undergo no change. If you worry that the witness could have lied, this should presumably  
 21 make you less confident about  $H=present$ . To accommodate this intuition, VERACITY can  
 22 be interpreted as a scenario in which an episode of awareness refinement takes place together  
 23 with a form of retraction. At first, after the learning episode, you update your belief based on  
 24 the *visual experience* of the witness. But after the growth in awareness, you realize that your  
 25 learning is in fact limited to what the witness *reported* to have seen. The previous learning  
 26 episode is retracted and replaced by a more careful statement of what you learned: instead  
 27 of conditioning on  $E=seen$ , you should condition on what the witness reported to have seen,  
 28  $R=seen-reported$ . This retraction will affect the probability of the hypothesis  $H=present$ .

29 Where does this leave us? Refinement cases that might at first appear similar can be  
 30 structurally different in important ways, and this difference can be appreciated by looking



1 at the Bayesian networks used to model them. In modeling VERACITY, the new node is  
2 added downstream, while in modeling LIGHTING, it is added upstream. This difference affects  
3 how probability assignments should be revised. Since the conditional probabilities associated  
4 with the upstream nodes are unaffected, Reverse Bayesianism is satisfied in VERACITY.<sup>9</sup> By  
5 contrast, since the conditional probabilities associated with the downstream node will often  
6 have to change, Reverse Bayesianism fails in LIGHTING.

7 This further corroborates our working hypothesis: structural features about how we con-  
8 ceptualize a specific scenario are the guiding principles about how we update the probability  
9 function through awareness growth, not a formal principle like Reverse Bayesianism. We  
10 further elaborate on this conjecture by drawing on some examples from Anna Mathani.

#### 11 4.1 Sure no-gain bets

12 Suppose the witness reports to have seeing the defendant around the crime scene. You are not  
13 aware that the witness could be lying. Thus, you are 100% confident that the witness saw is  
14 what they report to have seeing. In fact, you make no distinction between reporting to have  
15 seeing and seeing itself. So you would be willing to buy for 1\$ the following bet: if the witness  
16 saw the defendant, you get 1\$, and 0\$ otherwise. If the witness did see the defendant, you  
17 get you 1\$ back, and otherwise you loose \1\$. You are 100% sure the witness did see the  
18 defendant, so—by your lights—you stand to loose no money whatsoever from this bet. But  
19 suppose that, as a matter of fact, there is a difference between reporting and seeing. So, the  
20 witness might report to have seeing something without actually having seeing it. So, contrary  
21 to your conviction, that the witness saw the defendant is not 100% probable. This means that  
22 you would be willing to engage in a bet in which you are guaranteed not to win any money and  
23 could potentially lose money. If the witness did see the defendant you would get your 1\$ back,  
24 but if not, you would lose it.

### 25 5 Mathani's counterexamples

26 Mathani (2020) offers two counterexamples to Reverse Bayesianism. The first goes like this:

27 TENANT: Suppose that you are staying at Bob's flat which he shares with his

---

<sup>9</sup>Note that  $P(H=present|E=seen) \neq P(H=present|R=seen-reported)$ , but since you are conditioning on different propositions, this does not conflict with Reverse Bayesianism.

1 landlord. You know that Bob is a tenant, and that there is only one landlord, and  
 2 that this landlord also lives in the flat. In the morning you hear singing coming  
 3 from the shower room, and you try to work out from the sounds who the singer  
 4 could be. At this point you have two relevant propositions that you consider  
 5 possible ... *Landlord* standing for the possibility that the landlord is the singer,  
 6 and *Bob* standing for the possibility that Bob is the singer ... Because you know  
 7 that Bob is a tenant in the flat, you also have a credence in the proposition *Tenant*  
 8 that the singer is a tenant. Your credence in *Tenant* is the same as your credence  
 9 in *Bob*, for given your state of awareness these two propositions are equivalent  
 10 ... Now let's suppose the possibility suddenly occurs to you that there might be  
 11 another tenant living in the same flat (*Other*).

12 Initially, you thought the singer could either be the landlord or Bob, the tenant. Then you come  
 13 to the realization that a third person could be the singer, another tenant. Before awareness  
 14 growth, that Bob is in the shower and that a tenant is in the shower are equivalent descriptions.  
 15 After awareness growth, this equivalence breaks down.

16 Why is this scenario problematic? Suppose, after you hear singing in the shower, you become  
 17 sure someone is in there, but you cannot tell who. So  $P(\textit{Landlord}) = P(\textit{Bob}) = 1/2$ , and since  
 18 *Bob* and *Tenant* are equivalent, also  $P(\textit{Tenant}) = 1/2$ . Now, *Landlord*, *Bob* and *Tenant* are all  
 19 propositions that you were originally aware of, and thus Reverse Bayesianism requires that their  
 20 assigned probabilities should remain in the same proportion after your awareness grows. But  
 21 note that *Other* entails *Tenant* and *Bob* and *Other* are disjoint, so it follows that  $P^+(\textit{Other})$   
 22 must have zero probability.<sup>10</sup> This is an undesired outcome that rules out the possibility that  
 23 there could be a third person in the shower.<sup>11</sup>

24 Consider now Mathani's second counterexample:

25 COIN: You know that I am holding a fair ten pence UK coin which I am about to  
 26 toss. You have a credence of 0.5 that it will land *Heads*, and a credence of 0.5 that  
 27 it will land *Tails*. You think that the tails side always shows an engraving of a lion.  
 28 So you also have a credence of 0.5 that (*Lion*) it will land with the lion engraving  
 29 face-up: relative to your state of awareness *Tails* and *Lion* are equivalent.... Now

<sup>10</sup>If  $P^+(\textit{Other}) > 0$ , the proportion of *Tenant* to *Landlord* or the proportion of *Bob* to *Landlord* should change.

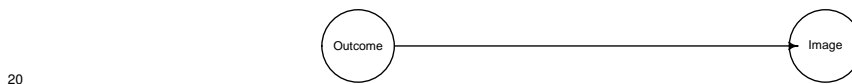
<sup>11</sup>Awareness Rigidity is no of help either because it would require that  $P^+(\textit{Landlord}|\textit{Landlord} \vee \textit{Tenant}) = P^+(\textit{Bob}|\textit{Landlord} \vee \textit{Tenant})$  both equal  $1/2$ , thus forcing  $P^+(\textit{Other}|\textit{Landlord} \vee \textit{Tenant})$  to zero.

1 let's suppose that you somehow become aware that occasionally ten pence coins  
 2 have .... an engraving of Stonehenge on the tails side.

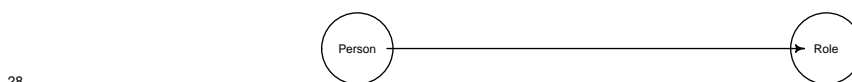
3 *Tails* and *Lion* are equivalent propositions prior to awareness growth. Suppose you initially  
 4 gave *Tails* and *Lion* the same credence. Reverse Bayesianism requires that their relative  
 5 proportions should stay the same after awareness grow. The same applies to *Heads* and *Tails*.  
 6 But since *Lion* and *Stonehenge* are incompatible and the latter entails *Tails*, you should have  
 7  $P^+(\text{Stonehenge}) = 0$ , again an undesirable conclusion.

8 Mathani notes that COIN has the same structure as TENANT. This is true to some extent, but  
 9 there is also an interesting asymmetry between the two scenarios. In TENANT, it is natural to  
 10 assign 1/3 to *Landlord*, *Bob* and *Other* after awareness growth. That someone is singing in  
 11 the shower is evidence that someone must be in there, but without any more discriminating  
 12 evidence, each person should be assigned the same probability. Consequently, a probability  
 13 of 2/3 should be assigned to *Tenant*. On this picture, the proportion of *Landlord* to *Tenant*  
 14 changes from 1:1 (before awareness growth) to 1:2 (after awareness growth). But, in COIN,  
 15 the relative proportion of *Heads* to *Tails* should remain constant throughout, unless evidence  
 16 emerges that the coin is not fair. One might have expected that *Landlord* and *Tenant* would  
 17 behave just like *Heads* and *Tails*, but actually they do not.

18 Bayesian networks can help to model the asymmetry between these two scenarios. Consider  
 19 COIN first. The structure of the scenario is represented by the following graph:



21 The upstream node *Outcome* has two states, *tails* and *heads*. These two states remain the same  
 22 throughout. What changes are the states associated with the *Image* node downstream. Before  
 23 awareness growth, the node *Image* has two states: *lions* and *heads-image*.<sup>12</sup> You assume that  
 24 *Image* = *lions* is true if and only if *Outcome* = *tails* is true. Then, you come to the realization  
 25 that the imagines for tails include a lion or a stonehenge engraving. So, after awareness growth,  
 26 the node *Image* contains three states: *lion*, *stonehenge* and *heads-image*. Consider now the  
 27 other scenario, TENANT. We start with the following graph:




---

<sup>12</sup>The heads side must have some image, not specified in the scenario.

1 Initially, the upstream node *Person* has two possible states, representing who is in the bathroom  
2 singing: *landlord-person* and *bob*. To simplify things, the assumption here is that the evidence  
3 of singing has already ruled out the possibility that no one would be in the shower. The  
4 downstream node *Role* has also two values, *landlord* and *tenant*. After your awareness grows,  
5 the upstream node *Person* should now have one more possible state, *other*.

6 The difference in modeling the two scenarios is this. In COIN, the states of the upstream  
7 node remain fixed, whereas in TENANT, they change. After awareness growth, no new state  
8 is added to *Outcome*, but an additional state, *other*, is added to *Person*. Plausible probability  
9 distributions for the Bayesian networks associated with the two scenarios are displayed in  
10 Table 1. How the networks should be built and which probabilities should shift is based on  
11 our background knowledge. This knowledge tells us that the equiprobability of *heads* and  
12 *tails* should not be affected by realizing that *stonhenge* is another possible engraving for the  
13 tails side. It also tells us that the probabilities of *landlord* and *tenant* should be affected by  
14 realizing that a third person could be in the shower.

15 We conclude with some programmatic remarks. We think that the awareness of agents grows  
16 while holding fixed certain material structural assumptions, based on commonsense, semantic  
17 stipulations or causal dependency.<sup>13</sup> To model awareness growth, we need a formalism that  
18 can express these material structural assumptions. This can be done using Bayesian networks,  
19 and we offered some illustrations of this strategy. These material assumptions also guide  
20 us in formulating the adequate conservative constraints, and these will inevitably vary on  
21 a case-by-case basis. The literature on awareness growth from a Bayesian perspective is  
22 primarily concerned with a formal, almost algorithmic solution to the problem. Insofar as  
23 Reverse Bayesianism is an expression of this formalistic aspiration, we agree with Steele and  
24 Stefánsson that we are better off looking elsewhere.

## 25 6 Towards a general theory

26 Awareness growth can occur in different ways. The key question is to what extent probability  
27 assignments that were made prior to the episode of awareness growth can be retained. There  
28 seems to be no clear rule that can decide that. We propose the following procedure. Construct a

---

<sup>13</sup> Arrows in Bayesian networks are often taken to represent causal relationships, but other interpretations exist. Schaffer (2016) discusses an interpretation in which arrows represent grounding relations rather than causality.

P(Image Outcome)		Outcome	
Image	lion	heads	tails
	heads-image	0	1
		1	0
P+(Image Outcome)		Outcome	
Image	lion	heads	tails
	stonehenge	0	1/2
	heads-image	0	1/2
		1	0
P(Outcome) = P+(Outcome)		Outcome	
	heads	tails	
	1/2	1/2	

P(Role Person)		Person		
Role	tenant	landlord-person	bob	
	landlord	0	1	
		1	0	
P+(Role Person)		Person		
Role	tenant	landlord-person	bob	other
	landlord	0	1/2	1/2
		1	0	0
P(Person)		Person		
	landlord-person	bob		
	1/2	1/2		
P+(Person)		Person		
	landlord-person	bob	other	
	1/3	1/3	1/3	

Table 1: Top table displays a plausible probability distribution for COIN and bottom table does the same for TENANT.

1 Bayesian network prior to awareness growth and compare it with the new Bayesian network  
2 after awareness growth. If the new arrows and nodes are all downstream, the old probabilities  
3 table should not be changed. The paradigmatic cases of this are scenarios VERACITY and  
4 COIN. If, instead, the new arrows and nodes are upstream, the old probabilities tables  
5 should be changed. The paradigmatic examples are LIGHTING and TENANT.

## 6 7 Counterexamples

7 In this section, we rehearse two of the counterexamples to Reverse Bayesianism by Steele and  
8 Stefánsson. One example targets awareness expansion and the other awareness refinement  
9 (more on this distinction soon). We show why they make a limited case against Reverse  
10 Bayesianism and then provide a better counterexample with the aid of Bayesian networks.

### 11 7.1 Friends and Movies

## 12 References

- 13 Bradley, R. (2017). *Decision theory with a human face*. Cambridge University Press.
- 14 Chihara, C. S. (1987). Some problems for bayesian confirmation theory. *British Journal for*  
15 *the Philosophy of Science*, 38(4), 551–560.
- 16 Earman, J. (1992). *Bayes or bust? A critical examination of bayesian confirmation theory*.  
17 MIT press.
- 18 Fenton, N., Neil, M., & Lagnado, D. A. (2013). A General Structure for Legal Arguments  
19 About Evidence Using Bayesian Networks. *Cognitive Science*, 37(1), 61–102. <https://doi.org/10.1111/cogs.12004>
- 20 Glymour, C. (1980). *Theory and evidence*. Princeton University Press.
- 21 Howson, C. (1976). The development of logical probability. In *Essays in memory of imre*  
22 *lakatos. Boston studies in the philosophy of science* (pp. 277–298). Springer.
- 23 Karni, E., & Vierø, M.-L. (2015). Probabilistic sophistication and reverse bayesianism. *Journal*  
24 *of Risk and Uncertainty Volume*, 50, 189–208.
- 25 Lakatos, I. (1968). Changes in the problem of inductive logic. *Studies in Logic and the*  
26 *Foundations of Mathematics*, 51, 315–417.
- 27 Mathani, A. (2020). Awareness growth and dispositional attitudes. *Synthese*, 198(9), 8981–  
28 8997.
- 29 Pettigrew, R. (forthcoming). How should your beliefs change when your awareness grows?  
30 *Episteme*, 1–25. <https://doi.org/10.1017/epi.2022.33>
- 31 Roussos, J. (2021). Awareness growth and belief revision. *Manuscript*.
- 32 Schaffer, J. (2016). Grounding in the image of causation. *Philosophical Studies*, 173, 49–100.
- 33 Steele, K., & Stefánsson, O. (2021). Belief revision for growing awareness. *Mind*, 130(520),  
34 1207–1232.
- 35 Wenmackers, S., & Romeijn, J.-W. (2016). New theory about old evidence: A framework for  
36 open-minded bayesianism. *Synthese Volume*, 193, 1225–1250.
- 37

- <sup>1</sup> Williamson, J. (2003). Bayesianism and language change. *Journal of Logic, Language, and*
- <sup>2</sup> *Information*, 12(1), 53–97.
- <sup>3</sup> Zabell, S. (1992). Predicting the unpredictable. *Synthese*, 90(1), 205–232.