

Awareness Growth with Bayesian Networks

no point in adding a date by hand if it's not systematically updated and confusing

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 and play a role in determining how probabilities should be updated in light of awareness growth. 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

1 specification of their ‘total possible future experience’ (Howson, 1976), as though learning was
2 confined to an ‘initial prison’ (Lakatos, 1968).

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

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

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

25 Let \mathcal{F} be the initial algebra of propositions and let \mathcal{F}^+ be the algebra after the agent’s
26 awareness state has grown. Both algebras contain the contradictory and tautologous propo-
27 sitions \perp and \top , and they are closed under connectives such as disjunction \vee , conjunction
28 \wedge and negation \neg . Denote by X and X^+ the subsets of these algebras that contain only basic
29 propositions, namely those without connectives. **Reverse Bayesianism** posits that the ratio

¹Roussos (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 of probabilities for any basic propositions A and B in both X and X^+ —the basic propositions
2 shared by the old and new algebra—remain constant through the process of awareness growth:

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

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

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

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

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

25 The plan for the paper is as follows. To set the stage for the discussion, Section 2 begins with
26 two counter-examples to Reverse Bayesianism by Steele & Stefánsson (2021). One example
27 targets awareness expansion and the other awareness refinement (more on the distinction soon).
28 Section 3 models cases of awareness expansion using Bayesian networks. Section 4 further

1 illustrates the fruitfulness of this approach by looking at two counter-examples to Reverse
2 Bayesianism by Mathani (2020). When modeled with Bayesian networks, Mathani's examples
3 are straightforward cases of awareness expansion. Section 5 turns to awareness refinement.
4 Finally, Section 6 outlines a general theory of awareness growth with Bayesian networks.

5 2 Steele and Stefánsson's Examples

6 In this section, we rehearse two of the counterexamples to Reverse Bayesianism by Steele &
7 Stefánsson (2021). One example targets awareness expansion and the other awareness refine-
8 ment. A precise definition of expansion can be tricky to provide, but a rough characterization
9 will suffice for now. Suppose, as is customary, propositions are interpreted as sets of possible
10 worlds, where the set of all possible worlds is the possibility space. Awareness expansion
11 occurs when a new proposition is added to the algebra and its interpretation includes possible
12 worlds not in the original possibility space. So the addition of the new proposition causes the
13 possibility space to expand. By contrast, awareness refinement (roughly) occurs when the new
14 proposition added to the algebra induces a more fine-grained partition of the possibility space.

15 The most straightforward case of *awareness expansion* occurs when you become aware
16 of a new explanation for the evidence at your disposal which you had not considered before.
17 This can happen in many fields of inquiry: medicine, law, science, everyday affairs. Here is a
18 scenario by Steele & Stefánsson (2021):

19 FRIENDS: Suppose you happen to see your partner enter your best friend's house
20 on an evening when your partner had told you she would have to work late. At
21 that point, you become convinced that your partner and best friend are having an
22 affair, as opposed to their being warm friends or mere acquaintances. You discuss
23 your suspicion with another friend of yours, who points out that perhaps they were
24 meeting to plan a surprise party to celebrate your upcoming birthday—a possibility
25 that you had not even entertained. (Steele & Stefánsson, 2021, sec. 5, Example 2)

26 Initially, the algebra contained the hypotheses 'my partner and my best friend met to have an
27 affair' (*affair*) and 'my partner and my best friend met as friends' (*friends*). These were the
28 only explanations you considered for the fact that your partner and your best friend met one
29 night without telling you. But, when the algebra changes, a new hypothesis is added which you

1 had not considered before: your partner and your best friends met to plan a surprise party for
2 your upcoming birthday (*surprise*).

3 This change in the algebra is not innocuous. At first, the hypothesis *affair* seems more likely
4 than *friends* because the former seems a better explanation than the latter for the secretive
5 behavior of your partner and best friend.² But when the new hypothesis *surprise* is added,
6 things change: *surprise* now seems a better explanation overall and thus more likely than
7 *affair*. And since *surprise* implies *friends*, the latter must be more likely than *affair*. This
8 conclusion violates Reverse Bayesianism since the ratio of the probabilities of *friends* and
9 *affair* has changed before and after awareness expansion:

$$\frac{P(\textit{affair})}{P(\textit{friends})} > 1 \text{ but } \frac{P^+(\textit{affair})}{P^+(\textit{friends})} < 1.$$

10 Steele & Stefánsson note that a quick fix is available. It is reasonable to suppose that no
11 change in the probabilities should occur so long as we confine ourselves to the old probability
12 space. With this in mind, consider the following condition, called **Awareness Rigidity**:

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

13 where T^* corresponds to a proposition that picks out, from the vantage point of the new
14 awareness state, the entire possibility space *before* the episode of awareness growth. Awareness
15 rigidity posits that, once a suitable proposition T^* is identified, the old probability assignments
16 remain unchanged conditional on T^* . In our running example, $\neg\textit{surprise}$ is the suitable propo-
17 sition T^* : *that* there were was no surprise party in the making picks out the original possibility
18 space. Conditional on $\neg\textit{surprise}$, no probability assignment should change, including the
19 probability of *affair*. This is the intended result.

20 But this is not the end of the story. Steele & Stefánsson go on to show that Awareness
21 Rigidity does not hold in other cases, what they call *awareness refinement*. As noted before,
22 these are cases in which the new proposition induces a more fine-grained partition of the
23 possibility space. Consider this scenario:

24 **MOVIES:** Suppose you are deciding whether to see a movie at your local cinema.

²This 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 with your best friend, the fact they behaved secretly or lied to you should not affect that much the probability of the two hypotheses.

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

11 You initially categorized movies by just language and genre, and then you refined your categorization by adding another variable, level of difficulty. Without considering language difficulty, you assigned the same probability to the hypotheses *thriller* and *comedy*. But learning that the owner was simple-minded made you think that the level of linguistic difficulty must be low and the movie most likely a thriller rather than a comedy (perhaps because thrillers are simpler—linguistically—than comedies). Since the probability of *thriller* goes up, this scenario violates (against Reverse Bayesianism) the condition $\frac{P(\textit{thriller})}{P(\textit{comedy})} = \frac{P^+(\textit{thriller})}{P^+(\textit{Comedy})}$. For the same reason, it also violates (against Awareness Rigidity) the condition $P(\textit{thriller}) = P^+(\textit{thriller}|\textit{thriller} \vee \textit{comedy})$, where $\textit{thriller} \vee \textit{comedy}$ is a proposition that picks out the entire possibility space.³

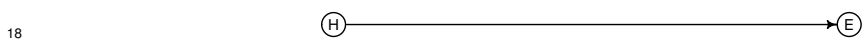
20 Some might object that the probability of *thriller* goes up, not because of awareness refinement, but because you learn that the owner is simple-minded. And if learning in the strict Bayesian sense—one modeled by conditionalization—takes place, it should be no surprise that some probabilities will change. Admittedly, the scenario can be split into two sub-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 new, namely that the owner is simple-minded. Arguably, the violation of Reverse Bayesianism or Awareness Rigidity can be attributed to learning that the owner is simple-minded, not to the mere growth in awareness. But there are cases of awareness refinement that do not involve learning in the Bayesian sense and still violate Reverse Bayesianism and Awareness Rigidity (see later in Section 5). So it is incumbent to understand under what circumstances these principles fail.

³Since MOVIES is a case of refinement, $\textit{thriller} \vee \textit{comedy}$ picks out the entire possibility space both before and after awareness growth.

1 As will become clear, we believe that theorizing about awareness growth should be grounded
2 in the subject-matter information underlying the scenario at hand. This subject-matter takes
3 many forms. In FRIENDS, awareness expansion does not change the basic presupposition that
4 someone’s behavior must have a reason. In MOVIES, awareness refinement does not change the
5 fact that characteristics such as language, difficulty or genre may influence one’s decision to
6 select a movie for showing rather than another. What is wrong with principles such as Reverse
7 Bayesianism or Awareness Rigidity—we hold—is that they are purely formal. We need instead
8 a formalism that can represent the relevant subject-matter information. In what follows, we
9 illustrate how Bayesian networks can serve this purpose. We first consider expansion (Section
10 3 and 4) and then refinement (Section 5).

11 3 Expansion with Bayesian Networks

12 A Bayesian network is a compact formalism to represent probabilistic dependencies. It consists
13 of a direct acyclic graph (DAG) accompanied by a probability distribution. The nodes in
14 the graph represent random variables that can take different values. We will use ‘nodes’ and
15 ‘variables’ interchangeably. The nodes are connected by arrows, but no loops are allowed,
16 hence the name ‘direct acyclic graph’. A simple graph structure we will use repeatedly is the
17 so-called *hypothesis-evidence idiom* (Fenton, Neil, & Lagnado, 2013):

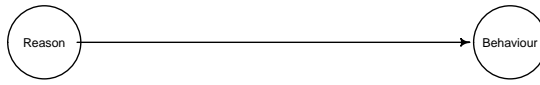


19 where H is the hypothesis node (upstream) and E the evidence node (downstream). If an
20 arrow goes from H to E , the full probability distribution associated with the Bayesian network
21 is defined by two probability tables.⁴ One table defines the prior probabilities $P(H = h)$ for
22 all the states (or values) of H , and another table defines the conditional probabilities of the
23 form $P(E = e|H = h)$, where uppercase letters represent the variables (nodes) and lower case
24 letters represent the states (or values) of these variables. These two probability tables are
25 sufficient to specify the full probability distribution. The other probabilities—say $P(E = e)$,
26 $P(H = h|E = e)$, etc.—follow by simply applying the probability axioms. As needed, more
27 complex graphical structures can also be used.

⁴A 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 18.

Bayesian networks are relied upon in many fields, but have been rarely deployed to model awareness growth (the exception is Williamson (2003)). We think instead they are a good framework for this purpose. Awareness growth can be modeled as a change in the graphical network—nodes and arrows are changed, added or erased—as well as a change in the probability distribution from the old to the new network. In this section, we focus on cases of awareness expansion and turn to refinement in the next section.

Recall the scenario FRIENDS from before. It can be modeled with the hypothesis-evidence idiom. The graph can be made more perspicuous by labeling the downstream node ‘Behavior’ (the evidence or fact to be explained) and the upstream node ‘Reason’ (the explanation or hypothesis about the the cause of the behavior):



Initially, before awareness growth, the hypothesis node *Reason* takes only two states: *friends* and *affair*. These two states are meant to be exclusive and exhaustive, so *affair* functions as the negation of *friends*, and vice versa. After awareness growth—specifically, awareness expansion—the two states are no longer exhaustive. A third state is added: *surprise*. So, in the formalism we are using, expansion simply consists in the addition of an extra state to one of the nodes of the network. The rest of the structure of the network remains intact.

Recall that the ratio of posterior probabilities of *friends* to *affairs* changed as a result of awareness expansion. The fact that your partner and best friend met without telling you—call this behavior *secretive*—initially made *affair* more likely than *friends*, but then the same fact made *friends* more likely than *affair* after awareness growth. So, formally,

$$\frac{P(\text{Reason}=\text{affair}|\text{Behavior}=\text{secretive})}{P(\text{Reason}=\text{friends}|\text{Behavior}=\text{secretive})} > 1 \text{ but } \frac{P^+(\text{Reason}=\text{friends}|\text{Behavior}=\text{secretive})}{P^+(\text{Reason}=\text{affair}|\text{Behavior}=\text{secretive})} > 1.$$

Despite this change in posterior probabilities, it is plausible to assume that the likelihoods do not change. Even after awareness expansion, the fact that your partner and best friend met without telling you—*secretive*—makes better sense in light of *affair* compared to *friends*:

$$\frac{P(\text{Behavior}=\text{secretive}|\text{Reason}=\text{affair})}{P(\text{Behavior}=\text{secretive}|\text{Reason}=\text{friends})} = \frac{P^+(\text{Behavior}=\text{secretive}|\text{Reason}=\text{affair})}{P^+(\text{Behavior}=\text{secretive}|\text{Reason}=\text{friends})} > 1.$$

This equality holds even though the novel explanation *surprise* introduced during awareness

1 expansion makes better sense of the secretive behavior overall:⁵

$$\frac{P^+(Behavior=secretive|Reason=surprise)}{P^+(Behavior=secretive|Reason=affair)} > 1.$$

2 This analysis suggests a slight reformulation of conditions such as Reverse Bayesianism
3 and Awareness Rigidity. Instead of comparing posterior probabilities of hypotheses given the
4 evidence ($P(H = h|E = e)$), we can compare which explanation or hypothesis makes better
5 sense of the evidence ($P(E = e|H = h)$). So, for all values h, h' and e of upstream node H and
6 downstream node E in the old network, consider the following constraint:

$$\frac{P(E = e|H = h)}{P(E = e|H = h')} = \frac{P^+(E = e|H = h \& X \neq x^*)}{P^+(E = e|H = h' \& X \neq x^*)}, \quad (C)$$

7 where x^* is the new state added and X is the node (upstream or downstream) to which the new
8 state belongs, such as $H = surprise$ in FRIENDS.

9 In words, the constraint says that one's old assessment of the relative plausibility of two
10 competing hypotheses in light of a fixed body of evidence should remain unchanged after
11 awareness expansion. Constraint (C) is a variant of Reverse Bayesianism that only applies to
12 conditional probabilities of the form $P(E = e|H = h)$ for Bayesian networks of the form $H \rightarrow E$.
13 In addition, the constraint mimics Awareness Rigidity in that it ensures that the conditional
14 probabilities in question exclude the novel state $X = x^*$.⁶

15 How generally does constraint (C) apply besides examples such as FRIENDS? We put
16 forward the following *working hypothesis*:

17 If there is no change in the structure of the network, constraint (C) holds generally.

18 If there is a change in the structure of the network, constraint (C) will fail under
19 certain conditions (to be specified later).

20 Since expansion—as we have defined it—does not change the network structure, the constraint
21 should always hold for expansion.⁷ We provide support for this claim in the next section.

⁵Even though $\frac{P^+(Behavior=secretive|Reason=surprise)}{P^+(Behavior=secretive|Reason=affair)} > 1$ and $\frac{P^+(Behavior=secretive|Reason=affair)}{P^+(Behavior=secretive|Reason=friends)} > 1$ —so *affair* still makes better sense of the evidence than *friends* before and after awareness expansion—the posterior probability of *friends* is higher than *affair* after awareness expansion.

⁶When the novel state is added to the upstream node H , the condition $X \neq x^*$ is redundant. We will see later cases in which the novel state is added to the downstream node E , and here the condition is not redundant.

⁷It is crucial that the new hypothesis or explanation does not change the existing structure of the network. Consider this example. You are wondering which horse will win the race. You have done a careful study of past performances under different conditions and concluded that Red is more likely to win than Green. But you

1 4 Mathani's Examples

2 To acquire a firmer grasp on constraint (C), we now examine a couple of examples by Mathani
3 (2020). These examples are intended as counterexamples to Reverse Bayesianism, as well as a
4 challenge to the distinction between expansion and refinement. When modeled with Bayesian
5 networks, however, Mathani's examples are straightforward cases of awareness expansion.

6 The first example goes like this:

7 TENANT: You are staying at Bob's flat which he shares with his landlord. You
8 know that Bob is a tenant, and that there is only one landlord, and that this landlord
9 also lives in the flat. In the morning you hear singing coming from the shower
10 room, and you try to work out from the sounds who the singer could be. At this
11 point you have two relevant propositions that you consider possible ... *landlord*
12 standing for the possibility that the landlord is the singer, and *bob* standing for
13 the possibility that Bob is the singer ... Because you know that Bob is a tenant
14 in the flat, you also have a credence in the proposition *tenant* that the singer is a
15 tenant. Your credence in *tenant* is the same as your credence in *bob*, for given
16 your state of awareness these two propositions are equivalent ... Now let's suppose
17 the possibility suddenly occurs to you that there might be another tenant living in
18 the same flat (*other*).

19 Initially, you thought the singer could either be the landlord or Bob, the tenant. Then you come
20 to the realization that a third person could be the singer, another tenant. The possibility that
21 there could be a third person in the shower—besides Bob or the landlord—is a novel explanation
22 for why you hear singing in the shower. So TENANT seems to be a standard case of expansion
23 like FRIENDS. At the same time, this scenario is a bit more complicated. The expansion in
24 awareness goes along with an interesting conceptual shift. Before awareness expansion, that
25 Bob is in the shower and that a tenant is in the shower are equivalent descriptions, but after the
26 expansion, this equivalence breaks down.

have not considered the possibility that Grey would run. If Grey does run, it will have a greater chance of winning than the others, but will also make—for some odd reason—Green a much better racer than Red. So the odds that Green will win compared to Red should now be higher. Here, the new hypothesis introduces novel information that was not known before, say that the participation of Grey would weaken Red's performance and strengthen Green's performance. So the network should be changed in two ways: first, a new state should be added to the outcome node (Green wins; Red wins; Grey wins); and second, a new node should be added modeling the fact that Grey is participating and its participation affects Red's and Green's performance.

1 As Mathani shows, this scenario challenges Reverse Bayesianism. For it is natural to assign
 2 $1/3$ to *landlord*, *bob* and *other* after awareness growth, and $1/2$ to *landlord* and *bob* before
 3 awareness growth. That someone is singing in the shower is evidence that someone must be in
 4 there, but without any more discriminating evidence, each person should be assigned the same
 5 probability. Consequently, a probability of $2/3$ should be assigned to *tenant* after awareness
 6 growth, but only $1/2$ before. On this picture, the proportion of *landlord* to *tenant* changes from
 7 1:1 (before awareness growth) to 1:2 (after awareness growth).^{8, 9}

8 One could resist the challenge. For recall that Reverse Bayesianism only applies to basic
 9 propositions, which we defined earlier as propositions without connectives. So a possible fix is
 10 to adopt the following principle: if two propositions happen to be equivalent relative to some
 11 awareness state, they cannot be both considered basic. In TENANT, since *bob* and *tenant* are
 12 initially equivalent descriptions of the same state of affairs, they would not be considered both
 13 basic propositions. If only *bob* is considered basic, along with *landlord*, then the proportion
 14 of the probability of *bob* and *landlord* would remain the same during awareness growth, but
 15 not the proportion of the probabilities of *tenant* and *landlord*. This yields the intuitive result.
 16 But it is odd that *landlord* would be considered basic, but not *tenant*. However, if *tenant* is
 17 considered basic along with *landlord*, Reverse Bayesianism would require that the proportion
 18 of the probability of *tenant* and *landlord* remain the same during awareness growth. This is
 19 counterintuitive.

20 So dividing propositions into basic and non-basic is riddled with difficulties. Still, the above
 21 discussion alerts us to the fact that a difference exists between propositions like *bob* and those
 22 like *tenant*. The latter describes a role that different people could play besides Bob. Bayesian
 23 networks can help to model the person/role distinction, as follows:



⁸Here is a more involved argument. Suppose, after you hear singing in the shower, you become sure someone is in there, but you cannot tell who. So $P(\text{landlord}) = P(\text{bob}) = 1/2$, and since *bob* and *tenant* are equivalent, also $P(\text{tenant}) = 1/2$. Now, *landlord*, *Bob* and *tenant* are all propositions that you were originally aware of, and thus Reverse Bayesianism requires that their probabilities should remain in the same proportion after your awareness grows. But note that *other* entails *tenant* and *bob* and *Other* are disjoint, so it follows that $P^+(\text{other})$ must have zero probability. If $P^+(\text{other})$ were greater than zero, the proportion of the probability of *tenant* to *landlord* (or the proportion of the probability of *bob* to *landlord*) should change.

⁹This scenario need not be a challenge for Awareness Rigidity. Much depends on the choice of the proposition T^* that picks out, from the vantage point of the new awareness state, the old possibility space prior to awareness growth. The proposition $\text{landlord} \vee \text{bob}$ does the job. For $P^+(\text{landlord}|\text{landlord} \vee \text{bob})$ and $P^+(\text{bob}|\text{landlord} \vee \text{bob})$ should both equal $1/2$, and thus $P^+(\text{other}|\text{landlord} \vee \text{bob}) = 0$, but this does not mean that $P^+(\text{other}|\text{landlord} \vee \text{tenant})$ should equal zero. This is the intended result.

$P(\text{Role} \text{Person})$		Person	
Role	tenant	landlord-person	bob
	landlord	0	1
	TOTAL	1	0
		1	1

$P^+(\text{Role} \text{Person})$		Person		
Role	tenant	landlord-person	bob	other
	landlord	0	1	1
		1	0	0
	TOTAL	1	1	1

$P(\text{Person})$		Person	
	landlord-person	bob	
	1/2	1/2	

$P^+(\text{Person})$		Person	
	landlord-person	bob	other
	1/3	1/3	1/3

Table 1: The table displays a plausible probability distribution for the TENANT scenarios. Constraint (C) is met.

1 This subject-matter information—the distinction between people and the role they play—remain
2 fixed throughout the process of awareness expansion. What changes is how the details are filled
3 in. Initially, the upstream node *Person* has two possible states, representing who could be in the
4 bathroom singing: *landlord-person* and *bob*.¹⁰ The downstream node *Role* has also two values,
5 *landlord* and *tenant*. After your awareness grows, the upstream node *Person* should now have
6 one more possible state, *other*. Crucially, note that since TENANT is a case of expansion by
7 our definition—a state was added to a node—constraint (C) should hold. This is precisely what
8 happens. After all, the conditional probabilities $P(\text{Role} = \text{landlord}|\text{Person} = \text{landlord})$ and
9 $P(\text{Role} = \text{landlord}|\text{Person} = \text{bob})$ do not change after awareness growth (see Table 1).

10 It is now time to test the tenability of constraint (C) in cases of awareness expansion a bit
11 further. So far we only considered cases of expansion in which a new state was added to an
12 *upstream* node. In FRIENDS, a state was added to the upstream node *Reason*, and in TENANT, a
13 state was added to the upstream node *Person*. What if the new state was added to a downstream
14 node? For consider a variation of FRIENDS. Suppose the downstream node *Behavior* could
15 initially take only two values, say *secretive* and *transparent*. That is, the behavior of your

¹⁰To simplify things, the assumption here is that the evidence of singing has already ruled out the possibility that no one would be in the shower. In principle, the network should be more complex and contain another node for the evidence to be explained (the fact of singing in the shower), as follows: $\text{Singing} \leftarrow \text{Person} \rightarrow \text{Role}$.

1 partner and best friend could be secretive or transparent. You then realize there could be
2 ambiguity in the behaviour, say the behavior could have elements of secrecy and elements
3 of transparency. So the node could take a third value: *ambiguous*. Initially, *secretive* and
4 *transparent* are considered exhaustive, but that is no longer true after adding *ambiguous*. So
5 the old conditional probabilities will change, specifically, $P(E = e|H = h) \neq P^+(E = e|H = h)$,
6 where H is the upstream node and E downstream. However, if we exclude the novel state, there
7 should be no change, so $P(E = e|H = h \& E \neq e^*) = P^+(E = e|H = h \& E \neq e^*)$, where e^* is
8 the novel state added to the downstream node E . Hence, constraint (C) should be satisfied.

9 The same analysis applies to a more complex example by Mathani:

10 COIN: You know that I am holding a fair ten pence UK coin which I am about
11 to toss. You have a credence of 0.5 that it will land *heads*, and a credence of 0.5
12 that it will land *tails*. You think that the tails side always shows an engraving of a
13 lion. So you also have a credence of 0.5 that it will land with the lion engraving
14 face-up (*lion*): relative to your state of awareness *tails* and *lion* are equivalent....
15 Now let's suppose that you somehow become aware that occasionally ten pence
16 coins have an engraving of Stonehenge on the tails side (*stonehenge*).

17 The propositions *tails* and *lion* are equivalent prior to awareness growth. Suppose you initially
18 gave *tails* and *lion* the same credence. If they are basic propositions, Reverse Bayesianism
19 would require that their relative proportions should stay the same after awareness grow. The
20 same is true of *heads* and *tails*. But since *lion* and *stonehenge* are incompatible and the latter
21 entails *tails*, you should have $P^+(\text{stonehenge}) = 0$, an undesirable conclusion.

22 Mathani observes that this scenario blurs the distinction between expansion and refinement.
23 For one thing, COIN seems a case of refinement. The space of possibilities is held fixed—the
24 coin could come up heads or tails—but the options for tails are further refined, for tails could
25 be *lion* or *stonehenge*. On the other hand, a new possibility has been added after awareness
26 growth, namely *stonehenge*, which had not been considered before. This would indicate that
27 COIN is a case of refinement. This ambiguity makes it difficult to settle whether the scenario is
28 a challenge for Awareness Rigidity.¹¹

¹¹If it is a case of refinement, $\text{heads} \vee \text{tails}$ would pick out the entire possibility space even before awareness growth. If so, by Awareness Rigidity, $P^+(\text{tails}|\text{heads} \vee \text{tails})$ and $P^+(\text{lion}|\text{heads} \vee \text{tails})$ should both equal $1/2$ since these were their probabilities before awareness growth. But these assignments would force $P^+(\text{stonehenge}|\text{heads} \vee \text{tails})$ to zero. To avoid this odd result for awareness Rigidity, one might argue that $\text{heads} \vee \text{tails}$ picks out a possibility space larger than the old one, because it also includes the possibility of

1 These difficulties disappear if the scenario is modeled using Bayesian networks. The
 2 definition of awareness expansion we have been working with is simple: whenever a new state
 3 is added to one of the nodes in the network, awareness expansion takes place. The novel state
 4 can be added to any node in the network. Each node, with its range of states, characterizes
 5 an exhaustive partition of the possibility space. Whenever a new state is added to a node, the
 6 partition associated with the node expands. (Refinement, as we will see in the next section,
 7 consists in the addition of a new node, not in the addition of a new state to an existing node.)

8 By the definition of expansion just given, COIN counts as a case of expansion, but it is
 9 structurally different from the more straightforward cases such as FRIENDS and TENANT.
 10 Bayesian networks can help to model the difference precisely. The scenario can be modeled by
 11 this familiar graph structure:



13 The upstream node *outcome* has two states, *tails* and *heads*. These two states remain the same
 14 throughout. What changes are the states associated with the *image* node downstream. Before
 15 awareness growth, the node *image* has two states: *lions* and *heads-image*.¹² You assume that
 16 *Image = lions* is true if and only if *Outcome = tails* is true. Then, you come to the realization
 17 that the imagines for tails could include a lion or a stonehenge engraving. So, after awareness
 18 growth, the node *Image* contains three states: *lion*, *stonehenge* and *heads-image*.

19 To some extent, COIN has the same structure as TENANT—they are modeled by the same
 20 networks structure—but there is also an important asymmetry that is apparent by comparing
 21 their Bayesian networks. In the network for COIN, the states of the upstream node remain fixed
 22 while a new state is added to the downstream node. In the network for TENANT, the opposite
 23 happens: the states of the downstream node remain fixed, while a new state is added to the
 24 upstream node. Specifically, after awareness expansion, no new state is added to upstream
 25 node *Outcome*, but an additional state, *other*, is added to the downstream node *Person*.

26 The states of the upstream node model what we view—loosely speaking—as more causally
 27 fundamental compared to the states of downstream nodes. In TENANT, that a person singing
 28 in the shower is more fundamental, while the proposition that describes the person’s roles is
 29 derivative, for multiple people could play the same role. In *Coin*, that an outcome (heads or

stonehenge. But arguably *heads* \vee *tails* should not pick out a larger possibility space. So which is it?

¹²The heads side must have some image, not specified in the scenario.

1 tails) is instantiated by different engravings is considered a derivative fact. What is causally
2 more fundamental is the outcome, not the engraving. Bayesian networks offer a language to
3 model these differences that are crucial to model episodes of awareness expansion.

4 So what about constraint (C)? It is easy to check that it is satisfied. Conditional probabil-
5 ities such as $P(\text{Image} = \text{lions} | \text{Outcome} = \text{tails})$ or $P(\text{Image} = \text{lions} | \text{Outcome} = \text{heads})$ remain
6 unchanged after awareness growth given the condition $\text{Image} \neq \text{stonehenge}$. Initially, $\text{Image} =$
7 lions is true if and only if $\text{Outcome} = \text{tails}$ is true. So, $P(\text{Image} = \text{lions} | \text{Outcome} = \text{tails})$
8 equals one, but it must also be that $P^+(\text{Image} = \text{lions} | \text{Outcome} = \text{tails} \ \& \ \text{Image} \neq \text{stonehenge})$
9 equals one. More generally, plausible probability distributions for the Bayesian networks
10 associated with the scenario COIN is displayed in Table 2. Constraint (C) is never violated.

11 All in all, examples in the literature that count as cases of expansion under our definition—
12 that is, a state is added to a network without changes in the network structure—obey constraint
13 (C). This provides good support for the first part of our working hypothesis: If there is no change
14 in the structure of the network, constraint (C) holds generally. In this sense, the constraint has
15 outperformed both Reverse Bayesianism and Awareness Rigidity. But our objective here is not
16 to replace one formal constraint with another. As noted in the introduction, we think that the
17 phenomenon of awareness growth in its generality cannot be modeled in a purely formal matter.
18 The success of constraint (C) relies on the right network structure. How the networks should be
19 built is based on our subject-matter knowledge—for example, that people’s behavior must have
20 a reason; that multiple peoples can play different roles; or that heads and tails can be associated
21 with different specific engravings. Constraint (C) holds when this subject-matter knowledge
22 does not change. However, sometimes awareness growth may bring in new subject-matter
23 knowledge and require changes to the structure of the network. This is our next topic.

24 5 Refinement with Bayesian Networks

25 To see how the network structure itself may require modifications, we turn now from cases of
26 expansion to cases of refinement. In the framework of Bayesian networks, expansion consists in
27 adding states to existing nodes in the network. Refinement, instead, can be modeled by adding
28 nodes to the network without adding any new state to existing nodes. Intuitively, refinement
29 takes place when an epistemic agent acquires a more-fined grained picture of the situation,

$P(\text{Image} \text{Outcome})$		<i>Outcome</i>	
<i>Image</i>	<i>lion</i>	<i>heads</i>	<i>tails</i>
	<i>heads-image</i>	0	1
		1	0
	TOTAL	1	1
$P^+(\text{Image} \text{Outcome})$		<i>Outcome</i>	
<i>Image</i>	<i>lion</i>	<i>heads</i>	<i>tails</i>
	<i>stonehenge</i>	0	1/2
	<i>heads-image</i>	0	1/2
		1	0
	TOTAL	1	1
$P(\text{Outcome}) = P^+(\text{Outcome})$		<i>Outcome</i>	
	<i>heads</i>	<i>tails</i>	
	1/2	1/2	

Table 2: Table displays a plausible probability distribution for the COIN scenario. Constraint (C) is met.

say instead of thinking that the political spectrum is divided into liberals and conservatives, the political spectrum can be further divided into traditional-liberal, new-liberal, traditional-conservative and new-conservative. The political spectrum is still divided into liberal and conservative—no expansion occurred—but the two categories have been further refined.¹³

Although there is no shortage of counterexamples to Reverse Bayesianism when it comes to awareness refinement, we will use our own. Recall that MOVIES—the refinement-based counterexample to Reverse Bayesianism by Steele & Stefánsson—suffered from a possible objection. The example contained awareness refinement paired with a standard case of Bayesian learning by conditionalization. Some might argue that conditionalization, not awareness refinement, is responsible for the change in probabilities. To alleviate this worry, we will work with our own example which can be more clearly interpreted as mere awareness refinement. Our own example will also allow us to underscore the role of subject-matter assumptions in theorizing about awareness growth. So consider this scenario:

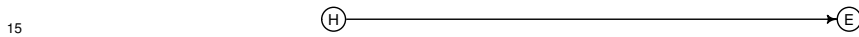
LIGHTING: You have evidence that favors a certain hypothesis, say a witness saw the defendant around the crime scene. You give some weight to this evidence. In your assessment, that the defendant was seen around the crime scene (your evidence) raises the probability that the defendant was actually there (your hypothesis). But now you ask, what if it was dark when the witness saw the defendant? In light of your realization that it could have been dark, you wonder whether (and

¹³This example is from Pettigrew (forthcoming).

1 if so how) you should change the probability that you assigned to the hypothesis
2 that the defendant was around the crime scene.

3 As your awareness grows, you do not learn anything specific about the lighting conditions,
4 neither that they were bad nor that they were good. You simply wonder what they were, a
5 variable you had previously not considered. So no Bayesian conditioning takes place in the strict
6 sense, although broadly speaking some new information has been introduced.¹⁴ Something has
7 changed in your epistemic state—you have a more fine-grained assessment of what could have
8 happened—but it is not clear what you should do in this scenario. Since the lighting conditions
9 could have been bad but could also have been good, perhaps you should just stay put until you
10 learn something more specific.

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



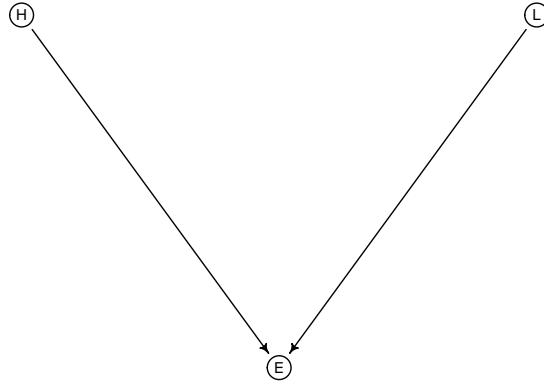
16 Since you trust the evidence, you think that the evidence is more likely under the hypothesis
17 that the defendant was present at the crime scene than under the alternative hypothesis:

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

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

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

¹⁴The 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 Bayesian updating, 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 of Bayesian updating.



1

2 where the node L can have two values, $L=good$ and $L=bad$. What is going on here? Initially,
3 you thought that the perceptual experience of the witness (node E) was causally affected by the
4 state of the whereabouts of the defendants (node H). But, as your awareness grows, you realize
5 that the witness' experience may also be caused by the environmental condition surrounding
6 the experience itself, say the lighting conditions (node L). So there are now two incoming
7 arrows into node E . In addition, commonsense as well as psychological findings suggest that
8 when the visibility deteriorates, people's ability to identify faces worsens. So a plausible way to
9 modify your assessment of the evidence is as follows:

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

10

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

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

15 Despite the change in awareness, you have not learned anything in the strict sense. Your
16 new stock of evidence does not contain neither the information that the lighting conditions
17 were bad nor that they were good. But the Bayesian network structure that represents your
18 epistemic state is now more fine-grained. The network contains the new variable L which it
19 did not contain prior to the episode of awareness growth. In addition—and this is the crucial
20 point—the new variable bears certain *structural relationships* with the variables H and E . The
21 graphical network represents a direct probabilistic dependency between the lighting conditions
22 L and the witness sensory experience E , but does not allow for any direct dependency between

1 the lighting conditions and the fact that the defendant was (or was not) at the crime scene.
 2 There is no direct arrow between the nodes L and H . This structure of dependencies captures
 3 our causal intuitions about the scenario: the lighting conditions do affect what the witness
 4 could see, but do not directly affect what the defendant might have done.

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

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

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

13 The question here is whether you should assess the evidence at your disposal—that the witness
 14 saw the defendant at the crime scene—any differently than before.¹⁵ As noted earlier, without
 15 a clear model of the scenario, it might seem that you should simply stay put. After all, besides
 16 the sensory experience of the witness, you have gained no novel information about the lighting
 17 conditions. Should you thus conclude that the evidence has the same value before and after the
 18 realization that lighting could have been bad?

19 The evidence would have the same value if the likelihood ratios associated with it relative to
 20 the competing hypotheses were the same before and after awareness growth. But, in changing
 21 the probability function from $P()$ to $P^+()$, it would be quite a coincidence if this were true. In
 22 our example, many possible probability assignments violate this equality. If before awareness
 23 growth you thought the evidence favored the hypothesis $H=\textit{present}$ to some extent, after the
 24 growth in awareness, the evidence is likely to appear less strong. To see this is tedious, so we
 25 relegate the details to a footnote.¹⁶ If this is correct, constraint (C) has been violated. Reverse

¹⁵Note that since no new state was added to an existing node, the condition $X \neq x^*$ in constraint (C) (where x^* is the new state added to an existing node X) is redundant here.

¹⁶By the law of total probability, the right hand side of the equality in (C) should be expanded, as follows:

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

- 1 Bayesianism is also violated since the ratio of the probabilities of $H=present$ to $E=seen$, before
 2 and after awareness growth, has changed:

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

- 3 where $P^{E=seen}()$ and $P^{+,E=seen}()$ represent the agent's degrees of belief, before and after aware-
 4 ness growth, updated by the evidence $E=seen$.¹⁷

- 5 The general lesson to be learned here has to do with the importance of formalizing subject-
 6 matter assumptions and the role of Bayesian networks in modeling awareness growth. Modeling
 7 the causal assumptions in LIGHTING allowed us to see that constraint (C)—as well as Reverse
 8 Bayesianism more generally—should fail here. But, as we noted already, constraint (C) might
 9 hold in other, structurally different cases of refinement. For consider this scenario:

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.

¹⁷The 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

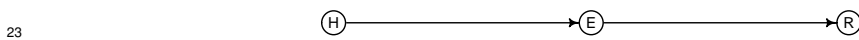
$$P^+(H=present|E=seen) \neq P(H=present|E=seen).$$

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

6 It might seem that this scenario is no different from LIGHTING. The realization that lighting
 7 could be bad should make you less confident in the truthfulness of the sensory evidence. And
 8 the same conclusion should presumably follow from the realization that the witness could be
 9 lying. But, upon closer scrutiny, things are not that simple. To run the two scenarios together
 10 would be a mistake.

11 The evidence at your disposal in LIGHTING is the sensory evidence—the experience of
 12 seeing—and the possibility of bad lighting does affect the quality of your visual experience.
 13 So, if lighting was indeed bad, this would warrant lowering your confidence in the truthfulness
 14 of the visual experience. But the possibility of lying in VERACITY does not affect the quality
 15 of the visual experience; it only affects the quality of the *reporting* of that experience. So, if
 16 the witness did lie, this would not warrant lowering your confidence in the truthfulness of the
 17 visual experience, only in the truthfulness of the reporting. The distinction between the visual
 18 experience and its reporting is crucial here. Bayesian networks help to model this distinction
 19 precisely, and then see why LIGHTING and VERACITY are structurally different.

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



24 As before, the hypothesis node H bears on the whereabouts of the defendant and has two values,
 25 $H=present$ and $H=absent$. Note the difference between E and R . The evidence node E bears
 26 on the visual experience had by the witness. The reporting node R , instead, bears on what the
 27 witness reports to have seen. The chain of transmission from ‘visual experience’ to ‘reporting’
 28 may fail for various reasons, such as lying or misremembering.

29 In VERACITY, the conditional probabilities, $P(E = e | H = h)$ should be the same as $P^+(E =$
 30 $e | H = h)$ for any values e and h of the variables H and E that are shared before and after

1 awareness growth. In comparing the old and new Bayesian network, this equality falls out from
2 their structure, as the connection between H and E remains unchanged. Thus, constraint (C) is
3 perfectly fine in scenarios such as VERACITY.

4 This does not mean that the assessment of the probability of the hypothesis $H=present$ should
5 undergo no change. If you worry that the witness could have lied, this should presumably
6 make you less confident about $H=present$. To accommodate this intuition, VERACITY can
7 be interpreted as a scenario in which an episode of awareness refinement takes place together
8 with a form of retraction. At first, after the learning episode, you update your belief based on
9 the *visual experience* of the witness. But after the growth in awareness, you realize that your
10 learning is in fact limited to what the witness *reported* to have seen. The previous learning
11 episode is retracted and replaced by a more careful statement of what you learned: instead
12 of conditioning on $E=seen$, you should condition on what the witness reported to have seen,
13 $R=seen-reported$. This retraction will affect the probability of the hypothesis $H=present$.

14 Where does this leave us? Refinement cases that might at first appear similar can be
15 structurally different in important ways, and this difference can be appreciated by looking at
16 the Bayesian networks used to model them. In modeling VERACITY, the new node is added
17 downstream, while in modeling LIGHTING, it is added upstream. This difference affects how
18 probability assignments should be revised. Since the conditional probabilities associated with
19 the upstream nodes are unaffected, constraint (C) is satisfied in VERACITY. By contrast, since
20 the conditional probabilities associated with the downstream node will often have to change,
21 the constraint fails in LIGHTING.

22 This further corroborates our working hypothesis: subject-matter assumptions—in the case
23 of LIGHTING and VERACITY, causal assumptions—about how we conceptualize a specific
24 scenario are the guiding principles about how we update the probability function through
25 awareness growth, not formal principles like Reverse Bayesianism, Awareness Rigidity or even
26 constraint (C).

27 **6 Towards a general theory**

28 We conclude with some programmatic remarks. We think that the awareness of agents grows
29 while holding fixed certain material structural assumptions, based on commonsense, semantic

1 stipulations or causal dependency.¹⁸ To model awareness growth, we need a formalism that
 2 can express these material structural assumptions. This can be done using Bayesian networks,
 3 and we offered some illustrations of this strategy. These material assumptions also guide
 4 us in formulating the adequate conservative constraints, and these will inevitably vary on
 5 a case-by-case basis. The literature on awareness growth from a Bayesian perspective is
 6 primarily concerned with a formal, almost algorithmic solution to the problem. Insofar as
 7 Reverse Bayesianism is an expression of this formalistic aspiration, we agree with Steele and
 8 Stefánsson that we are better off looking elsewhere.

9 Awareness growth can occur in different ways. The key question is to what extent probability
 10 assignments that were made prior to the episode of awareness growth can be retained. There
 11 seems to be no clear rule that can decide that. We propose the following procedure. Construct a
 12 Bayesian network prior to awareness growth and compare it with the new Bayesian network
 13 after awareness growth. If the new arrows and nodes are all downstream, the old probabilities
 14 table should not be changed. The paradigmatic cases of this are scenarios VERACITY and
 15 COIN. If, instead, the new arrows and nodes are upstream, the old probabilities tables
 16 should be changed. The paradigmatic examples are LIGHTING and TENANT.

17 References

- 18 Bradley, R. (2017). *Decision theory with a human face*. Cambridge University Press.
 19 Chihara, C. S. (1987). Some problems for bayesian confirmation theory. *British Journal for*
 20 *the Philosophy of Science*, 38(4), 551–560.
 21 Earman, J. (1992). *Bayes or bust? A critical examination of bayesian confirmation theory*.
 22 MIT press.
 23 Fenton, N., Neil, M., & Lagnado, D. A. (2013). A General Structure for Legal Arguments
 24 About Evidence Using Bayesian Networks. *Cognitive Science*, 37(1), 61–102. <https://doi.org/10.1111/cogs.12004>
 25 Glymour, C. (1980). *Theory and evidence*. Princeton University Press.
 26 Howson, C. (1976). The development of logical probability. In *Essays in memory of imre*
 27 *lakatos*. *Boston studies in the philosophy of science* (pp. 277–298). Springer.
 28 Karni, E., & Vierø, M.-L. (2015). Probabilistic sophistication and reverse bayesianism. *Journal*
 29 *of Risk and Uncertainty Volume*, 50, 189–208.
 30 Lakatos, I. (1968). Changes in the problem of inductive logic. *Studies in Logic and the*
 31 *Foundations of Mathematics*, 51, 315–417.
 32 Mathani, A. (2020). Awareness growth and dispositional attitudes. *Synthese*, 198(9), 8981–
 33 8997.
 34

¹⁸ 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.

- 1 Pettigrew, R. (forthcoming). How should your beliefs change when your awareness grows?
2 *Episteme*, 1–25. <https://doi.org/10.1017/epi.2022.33>
- 3 Roussos, J. (2021). Awareness growth and belief revision. *Manuscript*.
- 4 Schaffer, J. (2016). Grounding in the image of causation. *Philosophical Studies*, 173, 49–100.
- 5 Steele, K., & Stefánsson, O. (2021). Belief revision for growing awareness. *Mind*, 130(520),
6 1207–1232.
- 7 Wenmackers, S., & Romeijn, J.-W. (2016). New theory about old evidence: A framework for
8 open-minded bayesianism. *Synthese Volume*, 193, 1225–1250.
- 9 Williamson, J. (2003). Bayesianism and language change. *Journal of Logic, Language, and*
10 *Information*, 12(1), 53–97.
- 11 Zabell, S. (1992). Predicting the unpredictable. *Synthese*, 90(1), 205–232.